

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 10:40:04 ; Search time 36 Seconds
(without alignments)
1311.767 Million cell updates/sec

Title: US-09-589-870B-2

Perfect score: 936
Sequence: 1 MKRIVAAIAVSLTVSITA.....IDAKKAGVNGNPLDAVQO 183

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP rvirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #	Match Length	DB ID	Description
1	168.5	18.0	179	2	Q8KXW2 Rhizobium e
2	145.5	15.5	152	13	Q98SH4 Gallus gallus
3	114	12.2	529	5	Q25058 heliobacterium
4	112	12.0	595	2	Q06353 escherichia
5	110.5	11.8	882	2	Q9KX38 moraxella
6	105.5	11.3	1001	3	Q05164 saccharomyc
7	104	11.1	595	2	Q98077 escherichia
8	102	10.9	1310	16	Q8YJ44 listeria mo
9	102	10.9	1371	16	Q8XQ42 raietonia s
10	102	10.8	1374	16	Q8YJ44 bruceella me
11	101	10.8	595	2	Q06351 escherichia
12	100	10.7	560	2	Q984M2 escherichia
13	100	10.7	565	2	Q984M4 escherichia
14	99.5	10.6	1477	2	Q48028 haemophilus
15	99.5	10.6	2468	16	Q912M3 pseudomonas
16	99	10.6	1593	16	Q8FY73 bruceella su

17	98.5	10.5	832	2	O54356 moraxella c
18	98	10.5	595	2	O980U1 escherichia
19	98	10.5	595	16	Q8FGM2 escherichia
20	97.5	10.4	729	2	Q9KH34 antarctic b
21	97.5	10.4	3165	16	Q8XJ13 raietonia s
22	97.5	10.4	4106	16	Q8XOP2 raietonia s
23	97	10.4	443	6	Q28243 canis fam11
24	96	10.3	356	12	Q8V0L7 equine herp
25	96	10.3	389	12	Q8V0M0 equine herp
26	96	10.3	826	12	Q8V0L5 equine herp
27	96	10.3	4776	16	Q97P71 streptococ
28	95.5	10.2	744	3	Q8TFG9 schizosacch
29	95.5	10.2	881	2	Q9AF09 cellvibrio
30	95	10.1	122	13	Q8AY76 metagria g
31	94.5	10.1	438	16	Q8G3W5 streptococ
32	94.5	10.1	507	16	Q05056 streptococ
33	94.5	10.1	834	16	Q8DRT3 streptococ
34	94.5	10.1	3381	2	Q9KX33 pseudomonas
35	94	10.0	570	2	Q59665 pseudomonas
36	94	10.0	570	2	Q9AL29 escherichia
37	94	10.0	967	3	Q08294 saccharomyc
38	94	10.0	2190	16	Q8P6S0 xanthomonas
39	93.5	10.0	642	16	Q8XSW6 raietonia s
40	93.5	10.0	729	2	Q9EY54 vibrio sp.
41	93	9.9	1236	3	Q9C105 schizosacch
42	92.5	9.9	1018	16	Q9HNU6 pseudomonas
43	92.5	9.9	1113	16	Q8FXA7 bruceella su
44	92.5	9.9	1461	16	Q8XD19 escherichia
45	92	9.8	316	12	Q8V0M4 equine herp

ALIGNMENTS

RESULT 1
ID Q8KXW2 PRELIMINARY; PRT; 179 AA.
AC Q8KXW2; 01-OCT-2002 (TRENBLREL. 22, Created)
DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Hypothetical protein.
OS Rhizobium etli.
OC Plasmid symbiotic plasmid p42d.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RX MEDLINE=91193195; PubMed=2013564;
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
RT "Structural complexity of the symbiotic plasmid of Rhizobium
leguminosarum bv. phaseoli.";
RT J Bacteriol. 173:2411-2419(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RX MEDLINE=97419521; PubMed=9274036;
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
RT Cevallos M.A., Davila G.;
RT "Sequence, localization and characteristics of the replicator region
of the symbiotic plasmid of Rhizobium etli.";
RL Microbiology 143:2825-2831(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Quintero V., Cevallos M.A., Davila G.;
RT "A site-specific recombinase and RecA are required to exert
incompatibility towards the symbiotic plasmid of Rhizobium etli.";
RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80928; AAM5018.1; -

DT 01-MAR-2003 (TEMBLrel. 23, last annotation update)
 DE flagellin.
 GN flagellin.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B1 316-42 (Orskov);
 RA Schoenhaus G.J.;
 RL Thesis (1992), Unknown Institution.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B1 316-42 (Orskov);
 RX MEDLINE=93374833; PubMed=8366026;
 RA Schoenhaus G.J., Whitfield C.;
 RT "Comparative analysis of flagellin sequences from Escherichia coli
 RT strains possessing serologically distinct flagellar filaments with a
 RT shared complex surface pattern";
 RL J. Bacteriol. 175:5395-5402(1993).
 DR EMBL: L07388; AAA23799.1;
 DR InterPro: IPR001492; flagellin.
 DR InterPro: IPR001029; flagellin_C.
 DR Pfam: PF00700; flagellin_C.
 DR Pfam: PF00669; flagellin_N.
 DR PRINTS: PR00207; FLAGELLIN.
 DR PRODOM: PD000316; flagellin_C.
 DR SEQUENCE 595 AA; 6104 MW; 88510CD561EP25F9 CRC64;

Query Match 12.0%; Score 112; DB 2; Length 595;
 Best Local Similarity 26.1%; Pred. No. 3.9;
 Matches 52; Conservative 30; Mismatches 87; Indels 30; Gaps 9;

QY 7 AAIAVSLTVSITASADPSKQVSAEAGITGTYNQLGST--FIYTAGDGL 63
 ID 343 ASVTMGCTTYNFKTGADAD--AATNAGVSFTDTASKETVINKVATKQKAVAADDTSA 401
 DB 64 TGTYESAV-----GNAESRYVLTGRVDSAPAT---DGSGLAG-WIVAK--- 104
 DB 402 TTYKSGVQTVQAVFPAAGDSTASAKYADKADVSNATATYTDADGEMTTTGSYTTKYSIDA 461
 QY 105 NNYRNASHATTWGGQVVG--GAERINTQWLTS--GTTEANAKSTLVGH---FTKY 157
 DB 462 MNKVVVDSTGCTGKAPKAGVYVSANGTLTTDATSEGTVKDPLKALDEAISSIDKF 521
 QY 158 KPSAASIDAAKKAGVNNGN 176
 DB 522 RPSLGAIONRLDSAVTNLN 540

RESULT 5
 Q9KX38 PRELIMINARY; PRT; 892 AA.
 AC Q9KX38;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, last annotation update)
 DE USPAI.
 GN USPAI.
 OS Moraxella catarrhalis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Moraxella.
 OX NCBI_TaxID=480;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O46E;
 RX MEDLINE=2018164; PubMed=10671460;
 RA Lafontaine E.R., Cope L.D., Aebi C., Latimer J.L., McCracken G.H. Jr.,
 RA Hansen E.J.;
 RT "The Uspl1 Protein and a Second Type of Uspl2 Protein Mediate
 RT Adherence of Moraxella catarrhalis to Human Epithelial Cells in
 RT Vitro.";

RL J. Bacteriol. 182:1364-1373(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O46E;
 RA Maciver I., Latimer J.L., Cope L.D., Thomas S.E., Hansen E.J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U61725; AAF36416.1;
 DR InterPro: IPR000515; BPD.transp.
 DR InterPro: IPR005594; YadaK.
 DR Pfam: PF03895; YadaK.
 DR PROSITE: PS00402; BPD.TRANSF. INN MEMBR. 1.
 DR SEQUENCE 892 AA; 93358 MW; 9D71A36972F44C7 CRC64;

Query Match 11.8%; Score 110.5; DB 2; Length 892;
 Best Local Similarity 21.7%; Pred. No. 8;
 Matches 43; Conservative 26; Mismatches 86; Indels 43; Gaps 6;
 QY 17 SITASADPSKQVSAEAGITGTYNQLGSTFIYTAGDGLT--GTYESAVGNA 74
 DB 101 STVGGSSVTAKGEKSTIGGDTNDANGTYSTIGGYRAIGDSSTIGGTYNOATG-- 158
 QY 75 ESRVLTGRYDSAPATDGGTGALGWTAVKNNYR--NASHATTWGGQVVGAEARINTQ 131
 DB 159 EKSTVAGGRNQATGNNSTVAGGSYQATGNNSTVAGCSHNOATGCSFAAGVENKANN 218
 QY 132 -----WLTSTTEANAK-STLVGHDTFTYK 158
 DB 219 NVALGKNTTIDGNSVAIGSNNTTIDSGKQNVFIIGSSTNTTNAQGSVLLGHNTAGK-- 276
 QY 159 PSAASIDAAKKAGVNNGN 176
 DB 277 -KATAVSSAKVGLTLGN 293

RESULT 6
 ID 005164 PRELIMINARY; PRT; 1001 AA.
 AC 005164;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, last sequence update)
 DT 01-OCT-2002 (TEMBLrel. 22, last annotation update)
 DE AOB567, AOE1001, AOE110, AOE264 and AOE130 genes.
 GN AOE1001.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FY1679;
 RA Camo F.J., Lafuente M.J., Casamayor A., Aldea M., Casas C., Arrio J.,
 RA Herrero E., Gancedo C.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X89715; CAA61860.1;
 DR SEQUENCE 1001 AA; 97673 MW; F11BC4522DCFD84F CRC64;

Query Match 11.3%; Score 105.5; DB 3; Length 1001;
 Best Local Similarity 24.7%; Pred. No. 21;
 Matches 42; Conservative 33; Mismatches 64; Indels 31; Gaps 5;
 QY 6 VAAIAVSLTVSITASADPSKQVSAEAGITGTYNQLGSTFIYTAGDGLT--ADG 61
 DB 63 VSSSLTELTSSTEVSSIAFSTG---SEVSSITTSGSSVSGSSITTSGSSVSSS 118
 QY 62 ALTGTYESAVGNAESRYVLTGRYDSAPATDGGTGALGWTAVKNNYRNASHATTWGGQV 121
 DB 119 SATESGSSASGSSAT-----ESGSSVSGSSTSI-----TSSSSATSGSSSV 161
 QY 122 GAERINTQWLTTGTTTEANAKSTLVGHDTFTYKPSAASIDAAKKAG 171
 DB 162 SSSSTSI-----TSGSSATSGSSVSGSSTSGSSASGSSSSTATSG 205

```

RESULT 7
Q9S0T7 ID Q9S0T7 PRELIMINARY; PRT; 595 AA.
AC Q9S0T7;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Flagellin.
GN Flagellin.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bi 316-42;
RA Ohnishi K., Ishioka K., Matsuba T., Harayama S.;
RT "Cloning of H antigen genes in E.coli serotypes and expression in
RT E.coli K-12."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028475; BAAB5084.1;
DR InterPro; IPR001492; Flagellin_C.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR PRODOM; PD000316; Flagellin_C; 2.
SQ SEQUENCE 595 AA; 61020 MW; 8B9DC3D6ACT8C427 CRC64;

Query Match 11.1%; Score 104; DB 2; Length 595;
Best Local Similarity 25.6%; Pred. No. 14;
Matches 51; Conservative 30; Mismatches 88; Indels 30; Gaps 9;

QY 7 AAIASVLTITSASADPSKSKAQSAAEAGITGTWYNOLGST---FIYTAGADGAL 63
DB 343 ASVTMGCTTYNPFKGTADAD-AAATNANGVFTDTASKETLVNKVATKCKKAAALDGTISA 401
QY 64 TGTYESAV-----GMAESRYVLTGRYDSAPAT---DSGGTALG-VTWAMK--- 104
DB 402 TTTYSKGVOTYQAVFPAAGDGTASAKYADKADVSNATATYTDADGEMTTTIGSYTKYSIDA 461
QY 105 NNYRNASHATTWSGOVVG--GAEARINTQWLTS-GTTEANMKSTLVHD---TFTKV 157
DB 462 NNGKTVDSGTGTGKAPKVGAVVYANGTLTTPATSGVTYKDPKLAIDAISSIDKF 521
QY 158 KPSAASIDAKKAGVNGN 176
DB 522 RSSLGIHQRLDPSAVTNLN 540

RESULT 8
Q8Y4J4 ID Q8Y4J4 PRELIMINARY; PRT; 1310 AA.
AC Q8Y4J4;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein lmo2444.
GN LMO2444.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OC NCBI_TaxId=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Bagneri F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Cretouan F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussutger O.,
RA Entlan K.-D., Fshli H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Geobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz U., Kuhn M., Kunst F., Kwapack G.,

```

```

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
DR EMBL; AL591983; CAD00522.1;
DR Listlist; LMO2444;
DR InterPro; IPR005084; CBM 6.
DR InterPro; IPR00322; Glyco_hydro_31.
DR Pfam; PF03422; CBM_6; 1.
DR Pfam; PF01055; Glyco_hydro_31; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1310 AA; 144058 MW; D94919CA04021583 CRC64;

Query Match 10.9%; Score 102; DB 16; Length 1310;
Best Local Similarity 24.9%; Pred. No. 50;
Matches 48; Conservative 23; Mismatches 52; Indels 70; Gaps 9;

QY 12 SLTTSVITSASADPSKSKAQSAAEAGITGTWYNOLGSTFIYTAGADGAL-TGTYESA 70
DB 860 AVTYKVPQAQADAD-----TVITSGSASIVGTGYEME 893
QY 71 VGN-----AESRYVLTGRYDSAPATDGSCTALGVTWAMKN-----YRNAS- 112
DB 894 TGSNFPADTVADKPVAKETVY--DGYDK-----DGAGTIVANVKDSDGVNDLTYSKASSD 947
QY 113 ---ATTWSGOVYGAEARINTQWLTS-GTTEANMKSTL---VGHDTFKV----- 157
DB 948 NQALSTFYNGRYVQOTLKPTWDSVSETPLPSAGNKSISYKVTDTGKADQVSLDKV 1007
QY 158 ---KPSAASIDA 166
DB 1008 NIGFTPTVAKVEA 1020

RESULT 9
Q8XQ42 ID Q8XQ42 PRELIMINARY; PRT; 1371 AA.
AC Q8XQ42;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Putative hemagglutinin-related protein.
GN RSP1444 OR R803099.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OC NCBI_TaxId=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brotilier P., Camus J.C., Catalioco L.,
RA Chandier M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaepin C., Lavie M., Moisan A., Robert C., Savrin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646084; CAD18595.1;
DR InterPro; IPR006172; DNA_pol_B.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 1371 AA; 132558 MW; 1D94A3037BF8FC8F CRC64;

Query Match 10.9%; Score 102; DB 16; Length 1371;
Best Local Similarity 23.7%; Pred. No. 53;
Matches 52; Conservative 26; Mismatches 61; Indels 80; Gaps 9;

QY 8 AAIASVLTITSIT---SASADPSKSKAQS---AAEAGITGTWYNOLGSTFIYTA 57

```



```

Db 685 AILLTGATGATATMTWAGSGVTLALDLSLTSAGVATISGTSNSVSSRYRQGLSTVNIAGT 744
QY 58 GADGA-----LTGTYESAVGNAESRYVLTGRYDSAPATDG-----SGTA 96
Db 745 GSNASLITGCTVITGT-NSVGNASS-----TSNGNAVVLTDGKVLATATACPINIAGSN 799
QY 97 LGMTVAMKN-----NYRNAHSATTWAGGVGAERINTQWLTSCTT 139
Db 800 AGDGVWGSAGVMTSAPASSSITTSARSLDSVSGISGFYIGSGSKTL----- 848
QY 140 EANAAMKSTLVGHDFTFKVP---SAASIDAKKAGVNN 175
Db 849 -----TFATTAARVSLAIBSLVGARKAKFNNKG 874

RESULT 10
Q8YJ4M4 PRELIMINARY; PRT; 1374 AA.
ID O8YJ4M4;
AC O8YJ4M4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Extracellular serine protease (EC 3.4.21.-).
GN BMEI0058.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DeVecchio V.G., Kapral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lyfidis A., Reznik G.,
RA Jablonksi L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Lelsson J.-J.,
RA Haselkorn R., Kyriides N., Overbeek R.,
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL PROC. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AEO09449; AAL51240.1; -.
DR InterPro; IPR006315; Autotransporter.
DR InterPro; IPR005546; Autotransporter.
DR Pfam; PF03797; Autotransporter; 1.
DR TIGRPFAM; TIGR01414; autotrans_bar1. 2.
DR HydroLase; Complete proteome.
SQ SEQUENCE 1374 AA; 140577 MW; AE65C4BD5F60AB3 CRC64;

Query Match 10.8%; Score 102; DB 16; Length 1374;
Best Local Similarity 21.5%; Pred. No. 53;
Matches 45; Conservative 29; Mismatches 65; Indels 70; Gaps 7;

QY 20 ASASADPSKDSKAQVSAEAGI---TGTWYNOLGS-----TFIVTAGAD--- 60
Db 1078 APAGADPASSPVAGVITSDNGIWARIGDYSKLOSSSLTNMSONRTVITIGSGVCKFY 1137
QY 61 ----GALTGTYESAVGNAESRYVLTGRYDSAPATDGSCTALGW----- 99
Db 1138 EADTGKILGGINLALYSALISRI-----NSPFGDGPATTSAMGLGTLTWGSEGFYVD 1190
QY 100 ----TWAMKNYRNAHS-----ATTWAGGVGAERINTQWLTT-----S 136
Db 1191 GAOQINWYNNDYSDTAGKGLADKXATGYAVSIETGQRFNIGERWSVTPQAOQLMWSKLS 1250
QY 137 GTTEANAMKSTLVGHDFTFKVKSASID 165
Db 1251 MDTFNNIWEANVSLNDSLSLRAGVALD 1279

RESULT 11
Q06351 PRELIMINARY; PRT; 595 AA.
ID Q06351

```

```

AC 006351;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Flagellin.
GN FlIC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=Su 1242 (Orskov);
RA Schoenhalz G.J.;
RL Thesis (1992), Unknown Institution.
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=Su 1242 (Orskov);
RX MEDLINE=93374833; PubMed=8366026;
RA Schoenhalz G.J., Whitfield C.;
RT "Comparative analysis of flagellin sequences from Escherichia coli
RT strains possessing serologically distinct flagellar filaments with a
RT shared complex surface pattern."
RL J. Bacteriol. 175:5395-5402(1993).
DR EMBL; I07387; AAA23797.1; -.
DR InterPro; IPR001492; Flagellin.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_C; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C; 2.
SQ SEQUENCE 595 AA; 60923 MW; 9AAFB2E86884607A CRC64;

Query Match 10.8%; Score 101; DB 2; Length 595;
Best Local Similarity 25.6%; Pred. No. 24;
Matches 51; Conservative 29; Mismatches 89; Indels 30; Gaps 9;

QY 7 AAIANSLTIVSTTASASADPSKDSKAQVSAEAGITGTWYNOLGS---FIYTAGDGL 63
Db 343 ASVTWGGTGTYNFKTGADAG-AATNAGVSFTDTASKETVLNKAATKOGTAANAAGDTSA 401
QY 64 TGTYESAV-----GNAESRYVLTGRYDSAPAT---DGSCTALG-WTVAMK--- 104
Db 402 TTTTSGVOTYQAVPAAGGTASAKYADNTVSNATATYTDADGEMTTTIGSYTTKXSIDA 461
QY 105 NNRYRNAHSATTWAGGVG--GAERINTQWLTS-GTTEANAMKSTLVGHD---FTKY 157
Db 462 NNGKVTVDSGTSGKYAPVGAEVVVSANGTLTDTATSEGTVKDPLKALDEAIISSIDKF 521
QY 158 KPSASIDAKKAGVNNGN 176
Db 522 RSLSLGIONRLDSAVTNLN 540

RESULT 12
Q9S4M2 PRELIMINARY; PRT; 560 AA.
ID Q9S4M2;
AC Q9S4M2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Flagellin (Fragment).
GN FlIC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=E74/68;
RX MEDLINE=9084952; PubMed=9864325;
RA Reid S.D., Selander R.K., Whitam T.S.;
RT "Sequence diversity of flagellin (flc) alleles in pathogenic

```

RT Escherichia coli.;
RL U. Bacteriol. 181:153-160(1999).
DR EMBL; AF128954; AAD28525.2; -
DR InterPro; IPR001492; Flagellin_N.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_C; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR Prodom; PD000316; Flagellin_C; 2.
FT NON_TER 1
FT NON_TER 560
SQ SEQUENCE 560 AA; 57262 MW; 97DF005C9E0AB71 CRC64;

Query Match 10.7%; Score 100; DB 2; Length 560;
Best Local Similarity 25.1%; Pred. No. 26;
Matches 50; Conservative 30; Mismatches 89; Indels 30; Gaps 9;

QY 7 AAIIVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNOLGST---FIVTAGADAL 63
DB 324 ASVTMGITTYNFKTGADAD-AAITANAGVFTDTASKETVLNKVATAKGKAAADGDTSA 382
QY 64 TGTYESAV-----GNAESRYVLTGRYDSAPAT---DSSGTALG-WTVAMK--- 104
DB 383 TITVYSGVQTVQAVFAAGDGTASAKADADSNATATYTDADGEMTTIGSYTTKYSIDA 442
QY 105 NNRYRNAHSATTWSGOYV--GAERINTOWLLTS-GTTEANAMKSTLVGHD---TFPKV 157
DB 443 NNGKVTVDSGTGTGKAPKVAEYVVSANGTLTTDATSEGTVPKPLKALDEAISISIDKF 502
QY 158 KPSAASIDAKKAGVNNNGN 176
DB 503 RSSLGAIQNRLDSAVTNLN 521

RESULT 13
O9S4M4 PRELIMINARY; PRT; 565 AA.

AC O9S4M4; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Flagellin (Fragment).
GN FLIC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DEC 6a;
RX MEDLINE=99084952; PubMed=9864325;
RA Reid S.D., Selander R.K., Whitlam T.S.;
RT "Sequence diversity of flagellin (flc) alleles in pathogenic
RT Escherichia coli."
RL U. Bacteriol. 181:153-160(1999).
DR EMBL; AF128952; AAD28525.2; -
DR InterPro; IPR001492; Flagellin_N.
DR InterPro; IPR001029; Flagellin_C.
DR Pfam; PF00700; Flagellin_C; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR Prodom; PD000316; Flagellin_C; 2.
FT NON_TER 1
FT NON_TER 565
SQ SEQUENCE 565 AA; 57857 MW; 6E9E8A1AF72DC29 CRC64;

Query Match 10.7%; Score 100; DB 2; Length 565;
Best Local Similarity 25.1%; Pred. No. 26;
Matches 50; Conservative 30; Mismatches 89; Indels 30; Gaps 9;

QY 7 AAIIVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNOLGST---FIVTAGADAL 63
DB 324 ASVTMGITTYNFKTGADAD-AAITANAGVFTDTASKETVLNKVATAKGKAAADGDTSA 386

QY 64 TGTYESAV-----GNAESRYVLTGRYDSAPAT---DSSGTALG-WTVAMK--- 104
DB 387 TITVYSGVQTVQAVFAAGDGTASAKADADSNATATYTDADGEMTTIGSYTTKYSIDA 446
QY 105 NNRYRNAHSATTWSGOYV--GAERINTOWLLTS-GTTEANAMKSTLVGHD---TFPKV 157
DB 447 NNGKVTVDSGTGTGKAPKVAEYVVSANGTLTTDATSEGTVPKPLKALDEAISISIDKF 506
QY 158 KPSAASIDAKKAGVNNNGN 176
DB 507 RSSLGAIQNRLDSAVTNLN 525

RESULT 14
O48028 PRELIMINARY; PRT; 1477 AA.

AC O48028; 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Adhesin.
GN HMM2A.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12;
RX MEDLINE=92192797; PubMed=1548058;
RA Barenkamp S.U., Leininger E.;
RT "Cloning, expression, and DNA sequence analysis of genes encoding
RT nontypeable Haemophilus influenzae high-molecular-weight surface-
RT exposed proteins related to filamentous hemagglutinin of Bordetella
RT pertussis."
RL Infect. Immun. 60:1302-1313(1992).
DR EMBL; U08875; AAA20524.1; -
DR InterPro; IPR001000; Glyco_hydro.10.
DR InterPro; IPR000169; SHprot acate.
DR PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_His; 1.
SQ SEQUENCE 1477 AA; 154473 MW; B057C23F1AD24B0E CRC64;

Query Match 10.6%; Score 99.5; DB 2; Length 1477;
Best Local Similarity 26.2%; Pred. No. 87;
Matches 50; Conservative 26; Mismatches 88; Indels 27; Gaps 6;

QY 11 VSLTIVSITASADPSKDSKAQVSAEAGI---GTWYNOL-GSTFIVTAGADALTGT 66
DB 1177 ISGNTVSASAGDLTTRSGSKIEKSGEANTVATGTIGTISGNTVNVANA-GDLTVG 1235
QY 67 YESAVGNAESRYVLTGRYDSAPATDGGG-----TALGVVAMKNNRYRNAHSATTW 116
DB 1236 NGAIINATGGAATLTATGNTITTTAGSSITSTKQVULLAONGSIAGSINAANTLNTWG 1235
QY 117 SGQYVGAERINTOWLLTSGTTEANAMKSTLVGHDFTFKVKBPSA-----ASIDAAKA 170
DB 1296 TLTTVAGSDIK-----ATSGTLVINAKDAKLNDAAGDSTEVAANAAGSGSVTAATSS 1349
QY 171 GYNNGNPELDAY 181
DB 1350 SVNITGDLNTV 1360

RESULT 15
O912M3 PRELIMINARY; PRT; 2468 AA.

ID O912M3
AC O912M3; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein PA1874.

GN PA1874.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OK NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=2043737; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Gardner R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kae A., Larbig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004613; AG05263.1; .
 DR InterPro: IPR001343; Hemlysn_Ca_bind.
 DR InterPro: IPR006162; Ppantne_atlch.
 DR PRINTS; PRO0313; CABNDNGRPT.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 2468 AA; 238414 MW; 13596AFAB2C4B899 CRC64;

Query Match 10.6%; Score 99.5; DB 16; Length 2468;
 Best Local Similarity 24.9%; Pred. No. 1.6e+02;
 Matches 61; Conservative 19; Mismatches 70; Indels 95; Gaps 12;

QY 14 TTVSITASASA-----DPSKDSKAQVSAEAG-ITGTWYNQLGSTPIVAG--- 58
 DB 1381 TVNVNTASDAGNTSLPATTVDSSLPSIPQVDPNSGVSISGT--ADAGTIIITDNGN 1438
 QY 59 -----ADGA-----LTGTYESAVGNAESRYVLTGRYDSAPA----- 89
 DB 1439 PIGQVTVADSGNWSFTPGIPLPDGTVVNVVARSPS-----NVDSAPAVITVDGVAPAAP 1492
 QY 90 -----TDGSGTA-LGWTYAMKNNRNNAHSATTWGGVYGSEARINTQWLLTSGTEA 141
 DB 1493 VIDPSNGTEISGTAEAGATVI-----LTDGGGNPIGQATADSGNWTFTPTSTPLA 1542
 QY 142 NAWKSTLVGHDFFTKYK-PSAASIDA-----AKKAGV-----NNGN 176
 DB 1543 NGTVINAVADPAGNVTGPASVTVDAIAPPAPVINSNGVVISGTAAGATVILTDGNGN 1602
 QY 177 PLDAV 181
 DB 1603 PIGQV 1607

Search completed: October 27, 2003, 10:43:17
 Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993-- 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 10:40:04 ; Search time 11 Seconds

(without alignments)
782.353 Million cell updates/sec

Title: US-09-589-870B-2

Sequence: 1 MKRIVAAIAVSLTIVSITA.....IDAKKAGVNGNPLDVAQO 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	936	100.0	183	1 SAV_STRAV	P22629 streptomyc
2	932	99.6	183	1 SAVI_STRVL	O53532 streptomyc
3	897	95.8	183	1 SAV2_STRVL	O53533 streptomyc
4	146.5	15.7	150	1 AVR4_CHICK	P56734 gallus gall
5	143.5	15.3	152	1 AVID_CHICK	P02701 gallus gall
6	124.5	13.3	150	1 AVR6_CHICK	P56735 gallus gall
7	121.5	13.0	150	1 AVR3_CHICK	O13153 gallus gall
8	120.5	12.9	150	1 AVR7_CHICK	P56736 gallus gall
9	120.5	12.9	150	1 AVR3_CHICK	P56733 gallus gall
10	120	12.8	1064	1 FBPI_STRPU	P10079 stronglylce
11	116.5	12.4	150	1 AVR2_CHICK	P56732 gallus gall
12	100.5	10.7	1025	1 SLAP_CAVCR	P35828 cauliobacter
13	96.5	10.3	741	1 MAS2_MYCTU	O50596 mycobacteri
14	93	9.9	1225	1 Y309_MYCCE	P47551 mycoplasma
15	91.5	9.8	1130	1 STRP_ECOLI	P76072 escherichia
16	91.5	9.8	2249	1 OMPA_RICRI	P15921 rickettsia
17	90	9.6	1150	1 APMD_PIG	P12021 sus scrofa
18	89	9.5	255	1 CBPM_STRAL	P00733 streptomyc
19	89	9.5	303	1 HFAD_CAVCR	O45977 cauliobacter
20	89	9.5	488	1 PHB_ALCPA	P12625 alcaligenes
21	88.5	9.5	465	1 SLAP_LACBR	O05044 lacobacilli
22	87.5	9.3	457	1 PRIC_STRGR	P52320 streptomyc
23	87	9.3	666	1 MUR2_ENTHR	P39046 enterococcu
24	86.5	9.2	500	1 ABR1_TRIE	O92455 trichoderma
25	86	9.2	598	1 VAUF_SCHPO	O10168 schizosacch
26	85	9.1	797	1 VGLX_HSVBP	P28968 equine herp
27	85	9.1	872	1 GUNX_CELFI	P50401 cellulomona
28	84.5	9.0	962	1 GUNX_PSEFL	P10476 pseudomonas
29	83.5	8.9	500	1 XYLI_TRIKO	P48792 trichoderma
30	83.5	8.9	575	1 FLAB_CAMJE	P22251 campylobact
31	83.5	8.9	575	1 FLAB_CAMJE	P22252 campylobact
32	83.5	8.9	917	1 ICAS_MOUSE	O60655 mus musculu
33	83.5	8.9	1645	1 OMPB_RICTY	P96989 r outer mem

34	83.5	8.9	2021	1 OMPA_RICCN	O52657 rickettsia
35	83	8.9	610	1 CHIT_STRPL	P11220 streptomyc
36	82.5	8.8	297	1 PRTA_STRGR	P00776 streptomyc
37	82.5	8.8	331	1 ADT1_WHEAT	O41629 triticum ae
38	82.5	8.8	563	1 HEMA_IACKV	P09343 influenza a
39	82.5	8.8	616	1 MUTA_STRCM	O05064 streptomyc
40	82	8.8	381	1 ADT1_ARATH	P31167 arabidopsis
41	82	8.8	511	1 GUNB_PSEFL	P18126 pseudomonas
42	82	8.8	570	1 FBPI_STRPU	P49013 stronglylce
43	82	8.8	572	1 FLAB_CAMCO	P27053 campylobact
44	82	8.8	572	1 FLAB_CAMCO	P18245 campylobact
45	82	8.8	704	1 OE66_NPVAC	O00704 autocographa

ALIGNMENTS

RESULT 1
ID SAV_STRAV STANDARD: PRT; 183 AA.
AC P22629;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Streptavidin precursor.
OS Streptomyces avidinii.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Streptomycetales; Streptomycetaceae; Streptomycetes.
OX NCBI_Taxid=1895;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-64.
RX MEDLINE=86148514; PubMed=3951999;
RA Argarana C.E., Kuntz I.D., Birken S., Axel R., Cantor C.R.;
RT "Molecular cloning and nucleotide sequence of the streptavidin gene."; Nucleic Acids Res. 14:1871-1882(1986).
RN [2]
RP IMPORTANCE OF TRP IN BIOTIN-BINDING.
RX MEDLINE=89134083; PubMed=3223904;
RA Gitlin G., Bayer E.A., Wilchek M.;
RT "Studies on the biotin-binding site of streptavidin. Tryptophan residues involved in the active site."; Biochem. J. 256:279-282(1988).
RN [3]
RP IMPORTANCE OF TYR IN BIOTIN-BINDING.
RX MEDLINE=90351377; PubMed=2386489;
RA Gitlin G., Bayer E.A., Wilchek M.;
RT "Studies on the biotin-binding sites of avidin and streptavidin. Tyrosine residues are involved in the binding site."; Biochem. J. 269:527-530(1990).
RN [4]
RP CELL-BINDING.
RX MEDLINE=90358825; PubMed=2390089;
RA Alon R., Bayer E.A., Wilchek M.;
RT "Streptavidin contains an RVD sequence which mimics the RGD receptor domain of fibronectin."; Biochem. Biophys. Res. Commun. 170:1236-1241(1990).
RN [5]
RP X-RAY CRYSTALLOGRAPHY OF 37-157.
RX MEDLINE=97337436; PubMed=2911722;
RA Weber P.C., Orlowski D.H., Wendoloski J.J., Selenka F.R.;
RT "Structural origins of high-affinity biotin binding to streptavidin."; Science 243:85-88(1989).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 40-157.
RX MEDLINE=97337436; PubMed=2914176;
RA Freitag S., le Trong I., Klumb L., Steylen P.S., Stenkamp R.E.;
RT "Structural studies of the streptavidin binding loop."; Protein Sci. 6:1157-1166(1997).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.45 ANGSTROMS) OF 37-157.
RX MEDLINE=97294734; PubMed=9148939;
RA Katz B.A., Cases R.T.;
RT "In crystals of complexes of streptavidin with peptide ligands

RT containing the HPQ sequence the pKa of the peptide histidine is less
than 3.0.";
RT J. Biol. Chem. 272:13220-13228(1997).
RN [8]
RN X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 37-157.
RP MEDLINE=98070615; PubMed=9405158;
RX Katz B.A.;
RA "Binding of biotin to streptavidin stabilizes intersubunit salt
RT bridges between Asp61 and His87 at low pH.";
RL J. Mol. Biol. 274:776-800(1997).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 37-157 OF MUTANTS.
RX MEDLINE=98300340; PubMed=9636711;
RA Freitag S., Le Trong I., Chikoti A., Klumb L.A., Stayton P.S.,
RA Stenkamp R.E.;
RT "Structural studies of binding site tyrophophan mutants in the high-
RT affinity streptavidin-biotin complex.";
RL J. Mol. Biol. 279:211-221(1998).
CC -!- FUNCTION: THE BIOLOGICAL FUNCTION OF STREPTAVIDIN IS NOT KNOWN.
CC -!- FORMS A STRONG NON-COVALENT SPECIFIC COMPLEX WITH BIOTIN (ONE
CC MOLECULE OF BIOTIN PER SUBUNIT OF STREPTAVIDIN).
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
CC -!- DATABASE: NAME=Prozyme technical fact sheet;
CC WWW="http://www.prozyme.com/technical/sa10data.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X03591; CAA27265.1; -
DR PIR, A23513; A23513.
DR PDB, 1STP; 15-OCT-94.
DR PDB, 1PTS; 15-OCT-94.
DR PDB, 1SRE; 01-NOV-94.
DR PDB, 1SRF; 01-NOV-94.
DR PDB, 1SRG; 01-NOV-94.
DR PDB, 1SRH; 01-NOV-94.
DR PDB, 1SRH; 01-NOV-94.
DR PDB, 1SRI; 01-NOV-94.
DR PDB, 1SRJ; 01-NOV-94.
DR PDB, 1RST; 15-MAY-97.
DR PDB, 1RSU; 15-MAY-97.
DR PDB, 1STD; 03-APR-96.
DR PDB, 1SLE; 03-APR-96.
DR PDB, 1SLF; 03-APR-96.
DR PDB, 1SLG; 03-APR-96.
DR PDB, 1STR; 08-MAR-96.
DR PDB, 1STS; 08-MAR-96.
DR PDB, 1SWA; 04-MAR-98.
DR PDB, 1SWB; 04-MAR-98.
DR PDB, 1SWC; 04-MAR-98.
DR PDB, 1SWD; 04-MAR-98.
DR PDB, 1SWE; 04-MAR-98.
DR PDB, 1SWG; 15-JUL-98.
DR PDB, 1SMH; 23-MAR-99.
DR PDB, 1SMK; 23-MAR-99.
DR PDB, 1SMW; 23-MAR-99.
DR PDB, 1SMN; 23-MAR-99.
DR PDB, 1SMO; 23-MAR-99.
DR PDB, 1SWP; 23-MAR-99.
DR PDB, 1SMQ; 23-MAR-99.
DR PDB, 1SMR; 23-MAR-99.
DR PDB, 1VWA; 18-MAR-98.
DR PDB, 1VWC; 18-MAR-98.
DR PDB, 1VMD; 18-MAR-98.
DR PDB, 1VWE; 18-MAR-98.

DR PDB, 1VWF; 18-MAR-98.
DR PDB, 1VWG; 18-MAR-98.
DR PDB, 1VWH; 18-MAR-98.
DR PDB, 1VWI; 18-MAR-98.
DR PDB, 1VWJ; 18-MAR-98.
DR PDB, 1VWL; 18-MAR-98.
DR PDB, 1VWM; 18-MAR-98.
DR PDB, 1VMN; 18-MAR-98.
DR PDB, 1VMO; 18-MAR-98.
DR PDB, 1VMP; 18-MAR-98.
DR PDB, 1VMO; 18-MAR-98.
DR PDB, 1VMR; 18-MAR-98.
DR PDB, 1SMF; 29-APR-98.
DR PDB, 21ZA; 16-SEP-98.
DR PDB, 21ZB; 16-SEP-98.
DR PDB, 21ZC; 16-SEP-98.
DR PDB, 21ZD; 23-SEP-98.
DR PDB, 21ZE; 16-SEP-98.
DR PDB, 21ZF; 16-SEP-98.
DR PDB, 21ZG; 16-SEP-98.
DR PDB, 21ZH; 16-SEP-98.
DR PDB, 21ZI; 16-SEP-98.
DR PDB, 21ZJ; 16-SEP-98.
DR PDB, 21ZK; 16-SEP-98.
DR PDB, 21ZL; 16-SEP-98.
DR PDB, 2RTA; 18-NOV-98.
DR PDB, 2RTB; 18-NOV-98.
DR PDB, 2RTC; 18-NOV-98.
DR PDB, 2RTD; 18-NOV-98.
DR PDB, 2RTE; 18-NOV-98.
DR PDB, 2RTF; 27-JAN-99.
DR PDB, 2RTG; 18-NOV-98.
DR PDB, 2RTH; 18-NOV-98.
DR PDB, 2RTI; 25-NOV-98.
DR PDB, 2RTJ; 25-NOV-98.
DR PDB, 2RTK; 18-NOV-98.
DR PDB, 2RTL; 18-NOV-98.
DR PDB, 2RTM; 18-NOV-98.
DR PDB, 2RTN; 18-NOV-98.
DR PDB, 2RTO; 18-NOV-98.
DR PDB, 2RTP; 18-NOV-98.
DR PDB, 2RTQ; 18-NOV-98.
DR PDB, 2RTR; 18-NOV-98.
DR PDB, 1DF8; 20-SEP-00.
DR PDB, 1I9H; 25-DEC-02.
DR PDB, 1KEF; 01-MAY-02.
DR PDB, 1KL3; 01-MAY-02.
DR PDB, 1KL4; 01-MAY-02.
DR PDB, 1KL5; 01-MAY-02.
DR PDB, 1LCV; 06-NOV-02.
DR PDB, 1LCW; 06-NOV-02.
DR PDB, 1LCZ; 06-NOV-02.
DR PDB, 1SMS; 30-JUL-99.
DR PDB, 1SWT; 15-DEC-99.
DR PDB, 1SWU; 17-NOV-99.
DR InterPro: IPR005469; Avidin.
DR InterPro: IPR005468; Avidin/str.
DR InterPro: IPR005470; Streptavidin.
DR Pfam, PF01382; Avidin, 1.
DR PRINTS; PR00709; AVIDIN.
DR PROSITE; PS00577; AVIDIN; 1.
KW Signal; Biotin; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 183
FT BINDING 67 67
FT BINDING 78 78
FT BINDING 116 116
FT BINDING 132 132
FT BINDING 144 144
FT SITE 83 85
FT TURN 38 41
FT STRAND 43 47
STREPTAVIDIN.
INVOLVED IN BIOTIN BINDING.
INVOLVED IN BIOTIN BINDING.
INVOLVED IN BIOTIN BINDING.
INVOLVED IN BIOTIN BINDING.
INVOLVED IN BIOTIN BINDING.
CELL ATTACHMENT SITE.


```

FT BINDING 78 78 SIMILARITY) INVOLVED IN BIOTIN BINDING (BY
FT CHAIN 1 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 116 116 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 132 132 INVOLVED IN BIOTIN BINDING (BY
FT BINDING 144 144 INVOLVED IN BIOTIN BINDING (BY
SQ SEQUENCE 183 AA; 18833 MW; FEAFFDFE4ECCA CRC64;

Query Match
Best Local Similarity 95.8%; Score 897; DB 1; Length 183;
Pred. No. 1,5e-66;
Matches 174; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKRIVVAIAVSLTIVSITFASADPSKSKQVSAEAGITGTWYNOLGTFITAGAD 60
DB 1 MKRIVVAIAVSLTIVGITFASADPSKSKQVAAVAEAGITGTWYNOLGTFITANAD 60
QY 61 GALTGTYESAVGNASRYVLTGRYDAPATDGGTALGWTYAKNNYRNAHSATTWGGQY 120
DB 61 GALTGTYESAVGNASRYVLTGRYDAPATDGGTALGWTYAKNNYRNAHSATTWGGQY 120
QY 121 VGGARINTQWLLTSGTFEANAAMKSTLVGHDTFTKVKPSASIDAKKAGVNNGNPDA 180
DB 121 VGGARINTQWLLTSGTFEANAAMKSTLVGHDTFTKVKPSASIDAKKAGVNNGNPDA 180
QY 181 VQQ 183
DB 181 VQQ 183

RESULT 4
AVID CHICK STANDARD; PRT; 150 AA.
ID AV4 CHICK STANDARD; PRT; 150 AA.
AC P56734;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Avidin-related protein 4/5 precursor.
GN AV4 AND AV5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Oviduct;
RX MEDLINE=94170814; PubMed=8125122;
RA Keinaenen R.A., Wallen M.J., Kristo P.A., Laukkanen M.O.,
RA Toimela T.A., Helenius M.A., Kulomaa M.S.;
RT "Molecular cloning and nucleotide sequence of chicken avidin-related
RT genes 1-5.";
RL Eur. J. Biochem. 220:615-621(1994).
CC -1- MISCELLANEOUS: THE SEQUENCES OF THE CODING REGIONS OF GENES AV4
CC AND AV5 ARE IDENTICAL.
CC -1- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 222883; -; NOT_ANNOTATED_CDS.
CC PIR; S42204; S42204.
CC HSSP; P02701; 1RAV.
CC InterPro; IPR005468; Avidin/str.
CC Pfam; PF01382; Avidin; 1.
CC PROSITE; PS00577; AVIDIN; FALSE_NEG.

```

```

KM Biotin; Signal; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 150 AVIDIN-RELATED PROTEIN 4/5.
FT BINDING 57 57 BIOTIN (NON-COVALENT) (BY SIMILARITY).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16644 MW; 9A6C6C6310EFEL3A CRC64;

Query Match
Best Local Similarity 15.7%; Score 146.5; DB 1; Length 150;
Pred. No. 2.1e-05;
Matches 45; Conservative 22; Mismatches 47; Indels 19; Gaps 8;

QY 34 VSAEAGITGTWYNOLGTFITVA-GADGALTGTYESAV---GNAERYVLTGRYDSAP 88
DB 23 LSARKCSLTGKWTNNLGSIMTIRAVNSRGEFTGYTLTAVADPNPNTLSPULGIQHKRA- 81
QY 89 ATDGGTALGWTYAKNNYRNAHSATTWGGQY---GGAERINTQWLLTSGTFEAN-AW 144
DB 82 ---SQPTFGFTVW--NF--SESTTVFTGQCFIDRNGKEY-LKTMMLRSSVNDISYDW 132
QY 145 KSTLVGHDTFTKV 157
DB 133 KATRVGNNFTRL 145

RESULT 5
AVID CHICK STANDARD; PRT; 152 AA.
ID AV4 CHICK STANDARD; PRT; 152 AA.
AC P02701; 091958;
DT 21-OUL-1986 (Rel. 01, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Avidin precursor.
GN AVD.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87203384; PubMed=3575102;
RA Gope M.L., Keinaenen R.A., Kristo P.A., Conneely O.M., Beattie W.G.,
RA Zarucki-Schulz T., O'Malley B.W., Kulomaa M.S.;
RT "Molecular cloning of the chicken avidin cDNA.";
RL Nucleic Acids Res. 15:3595-3606(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90355928; PubMed=2143802;
RA Chandra G., Gray J.G.;
RT "Cloning and expression of avidin in Escherichia coli.";
RL Meth. Enzymol. 184:70-79(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Oviduct;
RX MEDLINE=95394357; PubMed=765080;
RA Wallen M.J., Laukkanen M.O., Kulomaa M.S.;
RT "Cloning and sequencing of the chicken egg-white avidin-encoding gene
RT and its relationship with the avidin-related genes Av1-Av5.";
RL Gene 161:205-209(1995).
RN [4]
RP SEQUENCE OF 25-152.
RX MEDLINE=71107558; PubMed=5100763;
RA Delange R.J., Huang T.-S.;
RT "Egg white avidin. 3. Sequence of the 78-residue middle cyanogen
RT bromide peptide. Complete amino acid sequence of the protein
RT subunit.";
RL J. Biol. Chem. 246:698-709(1971).
RN [5]
RP IMPORTANCE OF TYR IN BIOTIN-BINDING.
RX MEDLINE=90351377; PubMed=2386489;
RA Gilin G., Bayer E.A., Wilchek M.;

```


RT "Studies on the biotin-binding sites of avidin and streptavidin.
 RT Tyrosine residues are involved in the binding site.", streptavidin.
 RL Biochem. J. 269:527-530(1990).
 RN [6]
 RP BIOTIN-BINDING STUDIES.
 RX MEDLINE=91378911; Pubmed=1896347;
 RA Hiller Y., Bayer E.A., Milchek M.;
 RT "Studies on the biotin-binding site of avidin. Minimized fragments
 RT that bind biotin.";
 RL Biochem. J. 278:573-585(1991).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=93281699; Pubmed=8506353;
 RA Livnah O., Bayer E.A., Milchek M., Susman J.L.;
 RT "Three-dimensional structures of avidin and the avidin-biotin
 RT complex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:5076-5080(1993).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE=93294833; Pubmed=8515446;
 RA Pugliese L., Coda A., Malcovati M., Bolognesi M.;
 RT "Three-dimensional structure of the tetragonal crystal form of
 RT egg-white avidin in its functional complex with biotin at 2.7-A
 RT resolution.";
 RL J. Mol. Biol. 231:698-710(1993).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=98430987; Pubmed=9760187;
 RA Nardone E., Rosano C., Santambrogio P., Currin F., Corti A., Magni F.,
 RA Nicotrigio A.G., Paganeli G., Losso R., Aprea B., Bolognesi M.,
 RA Sidoli A., Arosio P.;
 RT "Biochemical characterization and crystal structure of a recombinant
 RT hen avidin and its acidic mutant expressed in *Escherichia coli*.";
 RL Eur. J. Biochem. 256:453-460(1998).
 CC -1- FUNCTION: THE BIOLOGICAL FUNCTION OF AVIDIN IS NOT KNOWN. FORMS A
 CC STRONG NON-COVALENT SPECIFIC COMPLEX WITH BIOTIN (ONE MOLECULE OF
 CC BIOTIN PER SUBUNIT OF AVIDIN).
 CC -1- SUBUNIT: Homotetramer.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN HEN OVIDUCT AND CONCENTRATED IN
 CC EGG WHITE (WHERE IT REPRESENT 0.05% OF THE TOTAL PROTEIN).
 CC -1- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
 CC -1- DATABASE: NAME=Worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/AV.html".
 CC -1- DATABASE: NAME=Prozyme technical fact sheet;
 CC WWW="http://www.prozyme.com/technical/av10data.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X05343; CAA28954.1; -
 CC EMBL: L27818; AAB59733.1; -
 CC PIR: A54975; VICH.
 DR PDB: 2AVI; 15-VUL-93.
 DR PDB: 1AVD; 31-JAN-94.
 DR PDB: 1AVE; 31-JAN-94.
 DR PDB: 1RAV; 15-JUL-98.
 DR PDB: 2CAM; 15-JUL-98.
 DR PDB: 1138; 25-DEC-02.
 DR PDB: 1LDO; 06-NOV-02.
 DR PDB: 1LDQ; 06-NOV-02.
 DR PDB: 1LEL; 06-NOV-02.
 DR InterPro: IPR005469; Avidin.
 DR InterPro: IPR005468; Avidin.str.
 DR Pfam: PF01382; Avidin.1.
 DR PRINTS: PR00709; AVIDIN.
 DR PROSITE: PS00577; AVIDIN.1.
 DR GlycoProtex: Signal; Biotin; 3D-structure.
 KW SIGNAL 1 24

FT CHAIN 25 152 AVIDIN.
 FT DISULFID 28 107 N-LINKED (GLUCNAC.).
 FT CARBOHYD 41 41 BIOTIN (NON-COVALENT).
 FT BINDING 57 57 I -> T (IN APPROX. 50% OF THE CHAINS).
 FT VARIANT 58 58 G -> S (IN REF. 3).
 FT CONFLICT 22 22 E -> Q (IN REF. 2 AND 3).
 FT CONFLICT 77 77
 FT STRAND 32 36
 FT STRAND 41 44
 FT TURN 49 50
 FT STRAND 52 59
 FT STRAND 69 77
 FT HELIX 80 82
 FT STRAND 87 92
 FT STRAND 100 109
 FT STRAND 115 124
 FT HELIX 130 135
 FT STRAND 137 146
 SQ SEQUENCE 152 AA; 16769 MW; 1D55A4491D5EPD5C CRC64;
 Query Match. 15.3%; Score 143.5; DB 1; Length 152;
 Best Local Similarity 33.3%; Pred. No. 3.8e-05;
 Matches 44; Conservative 19; Mismatches 56; Indels 13; Gaps 7;
 QY 34 VSAAEAGINGTWNQSGTFIVTA-GADGALGTYESAVGNMESRYVLTCRYDSAPATDG 92
 DB 23 LSRKCSLTGKNTNDGSMNTTGAVNSRGEFTGYTAV-TATSNEIKESPLHGTENTIN 81
 QY 93 SGT--ALGVTVMKNNVRAHSATTWGGQV---GGAEARINTQWLLTSGTTE-ANAMKS 146
 DB 82 KRTQPFPGFTVMK-----FSESTVFTGTCFLIDRCKEY-LKTMWLLRSSVNDIGDWMKA 136
 QY 147 TLVGHDTFTKVK 158
 DB 137 TRVGINIFTRLR 148
 RESULT 6
 ID_AVR6_CHICK STANDARD; PRT; 150 AA.
 AC P56735;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Avidin-related protein 6 precursor.
 GN AVR6.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rhode Island;
 RX MEDLINE=21070478; Pubmed=11167523;
 RA Ahlroth M.K., Kola E.H., Ewald D., Maaabanda J., Sazanov A., Pries R.,
 RA Kulomaa M.S.;
 RT "Characterization and chromosomal localization of the chicken avidin
 RT gene family.";
 RL Anim. Genet. 31:367-375(2000).
 CC -1- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ237658; CAB39893.1; -
 CC DR HSSP: P02701; 1RAV.
 DR InterPro: IPR005468; Avidin.str.

```

DR Pfam; PF01382; Avidin; 1.
DR PROSITE; PS00577; AVIDIN; 1.
KW Biotin; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 150
FT BINDING 57 57 BIOTIN (NON-COVALENT) (BY SIMILARITY).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16528 MW; CQF357D6291B150 CRC64;

Query Match
Best Local Similarity 32.8%; Score 124.5; DB 1; Length 150;
Matches 44; Conservative 25; Mismatches 42; Indels 23; Gaps 10;

QY 34 VSAEAGITGTWYNOLGSTFIYTAGA---DGLTGYESAV---GNAESRYVLTGRYDS 86
DB 23 LSARKCSLTGKMDNDLGS--IMTIGAVNDNGEFNGTYITAVADNPGNITRSPILGIQHKR 80
QY 87 APATDGSCTALGWTVMKNNYRNHSAITWSGQ-YV--CGAEARINTOMLTSGTTE-AN 142
DB 81 A-----CQPTFGFTVHM--NF--SESTSVFVGQCFFVDRSGREV-LKTKWLQRLAVDDISD 130
QY 143 AMKSTLVGHDTFTK 156
DB 131 DMKATRVGNDFTR 144

RESULT 7
AVR1_CHICK
ID AVR1_CHICK STANDARD; PRT; 150 AA.
AC O13153;
DT 15-JUL-1999 (Rel. 38, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Avidin-related protein 1 precursor.
GN AVR1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Oviduct;
RA MEDLINE=88260103; PubMed=2838690;
RA Keinaenen R.A., Laukkanen M.-L., Kulomaa M.S.;
RT "Molecular cloning of three structurally related genes for chicken
RT avidin."
RL J. Steroid Biochem. 30:17-21 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Oviduct;
RX MEDLINE=94170814; PubMed=8125122;
RA Keinaenen R.A., Wallen M.J., Kristo P.A., Laukkanen M.O.,
RA Toimela T.A., Helenius M.A., Kulomaa M.S.;
RT "Molecular cloning and nucleotide sequence of chicken avidin-related
RT genes 1-5."
RL Eur. J. Biochem. 220:615-621 (1994).
CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z21611; -; NOT ANNOTATED_CDS.
DR EMBL; Z97063; CAB09758.1; -
DR PIR; S42201; S42201.
DR HSP; P02701; IRAV.

```

```

DR InterPro; IPR005469; Avidin.
DR InterPro; IPR005468; Avidin/str.
DR Pfam; PF01382; Avidin; 1.
DR PRINTS; PR00709; AVIDIN.
DR PROSITE; PS00577; AVIDIN; 1.
KW Biotin; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 150
FT BINDING 57 57 BIOTIN (NON-COVALENT) (BY SIMILARITY).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16468 MW; 77548380B95B1f6 CRC64;

Query Match
Best Local Similarity 32.8%; Score 121.5; DB 1; Length 150;
Matches 44; Conservative 24; Mismatches 43; Indels 23; Gaps 10;

QY 34 VSAEAGITGTWYNOLGSTFIYTAGA---DGLTGYESAV---GNAESRYVLTGRYDS 86
DB 23 LSARKCSLTGKMDNDLGS--IMTIGAVNDNGEFNGTYITAVADNPGNITRSPILGIQHKR 80
QY 87 APATDGSCTALGWTVMKNNYRNHSAITWSGQ-YV--CGAEARINTOMLTSGTTE-AN 142
DB 81 A-----CQPTFGFTVHM--NF--SESTSVFVGQCFFVDRSGREV-LKTKWLQRLAVDDISD 130
QY 143 AMKSTLVGHDTFTK 156
DB 131 DMKATRVGNDFTR 144

RESULT 8
AVR7_CHICK
ID AVR7_CHICK STANDARD; PRT; 150 AA.
AC P56736; O9M6V4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Avidin-related protein 7 precursor.
GN AVR7.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Rhode Island.
RX MEDLINE=21070478; PubMed=1167523;
RA Ahlroth M.K., Kola E.H., Ewald D., Masabanda J., Sazanov A., Fries R.,
RA Kulomaa M.S.;
RT "Characterization and chromosomal localization of the chicken avidin
RT gene family."
RL Anim. Genet. 31:367-375 (2000).
CC -!- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ237659; CAB39894.1; -
DR HSP; P02701; IRAV.
DR InterPro; IPR005468; Avidin/str.
DR Pfam; PF01382; Avidin; 1.
DR PROSITE; PS00577; AVIDIN; 1.
KW Biotin; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 150
FT BINDING 57 57 BIOTIN (NON-COVALENT) (BY SIMILARITY).

```


CC DURING EARLY CLEAVAGE, THEN RAPIDLY INCREASES IN ABUNDANCE BETWEEN
CC LATE MORULA AND MESENCHYME BLASTULA STAGES TO MAXIMAL LEVELS
CC MAINTAINED THROUGH SUBSEQUENT STAGES. EXPRESSED BOTH MATERNALLY
CC AND ZYGOTICALLY.
CC -1- SIMILARITY: Contains 21 EGF-like domains.
CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR TO
CC AVIDIN/STREPTAVIDIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, L08692; AAA62164.1; -
CC EMBL, L08692; AAA62163.1; -
CC EMBL, X17530; CAA35571.1; -
CC EMBL, M17421; AAA30050.1; -
CC EMBL, X17533; CAA35573.1; -
CC PIR, A40136; A40136.
CC HSP, P01132, 1EGF.
CC DR InterPro, IPR000152; Asx hydroxyl.
CC DR InterPro, IPR005469; Avidin.
CC DR InterPro, IPR005468; Avidin/str.
CC DR InterPro, IPR000859; CUB domain.
CC DR InterPro, IPR000742; EGF_2.
CC DR InterPro, IPR001881; EGF_Ca.
CC DR InterPro, IPR001438; EGF_11.
CC DR Pfam, PF01382; Avidin; 1.
CC DR Pfam, PF01382; Avidin; 1.
CC DR Pfam, PF00431; CUB; 1.
CC DR Pfam, PF00008; EGF; 21.
CC PRINTS; PR00709; AVIDIN.
CC DR SMART; SMO0010; EGFBLD.
DR SMART; SMO0042; CUB; 1.
DR PROSITE, PS000179; EGF CA; 20.
DR PROSITE, PS000022; EGF_1; 19.
DR PROSITE, PS00577; AVIDIN; 1.
DR PROSITE, PS01180; CUB; 1.
DR PROSITE, PS01186; EGF_2; 19.
DR PROSITE, PS01187; EGF CA; 18.
KW Biocin: Alternative splicing: EGF-like domain; Repeat; Signal;
KW Glycoprotein; Calcium-binding
FT SIGNAL 1 19
FT CHAIN 20 1064
FT DOMAIN 20 55
FT DOMAIN 62 175
FT DOMAIN 176 212
FT DOMAIN 214 250
FT DOMAIN 252 288
FT DOMAIN 290 326
FT DOMAIN 328 364
FT DOMAIN 366 402
FT DOMAIN 404 440
FT DOMAIN 442 478
FT DOMAIN 480 516
FT DOMAIN 518 554
FT DOMAIN 556 592
FT DOMAIN 594 630
FT DOMAIN 632 668
FT DOMAIN 670 706
FT DOMAIN 708 744
FT DOMAIN 746 782
FT DOMAIN 784 820
FT DOMAIN 822 858
FT DOMAIN 860 896
FT DOMAIN 898 934
FT DOMAIN 936 1064
FT DISULFID 23 34

FT DISULFID 28 43 BY SIMILARITY.
FT DISULFID 45 54 BY SIMILARITY.
FT DISULFID 180 191 BY SIMILARITY.
FT DISULFID 185 200 BY SIMILARITY.
FT DISULFID 202 211 BY SIMILARITY.
FT DISULFID 218 229 BY SIMILARITY.
FT DISULFID 223 238 BY SIMILARITY.
FT DISULFID 228 249 BY SIMILARITY.
FT DISULFID 240 256 BY SIMILARITY.
FT DISULFID 256 267 BY SIMILARITY.
FT DISULFID 261 276 BY SIMILARITY.
FT DISULFID 278 287 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 316 325 BY SIMILARITY.
FT DISULFID 332 343 BY SIMILARITY.
FT DISULFID 337 352 BY SIMILARITY.
FT DISULFID 354 363 BY SIMILARITY.
FT DISULFID 370 381 BY SIMILARITY.
FT DISULFID 375 390 BY SIMILARITY.
FT DISULFID 392 401 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 428 BY SIMILARITY.
FT DISULFID 430 439 BY SIMILARITY.
FT DISULFID 446 457 BY SIMILARITY.
FT DISULFID 451 466 BY SIMILARITY.
FT DISULFID 468 477 BY SIMILARITY.
FT DISULFID 484 495 BY SIMILARITY.
FT DISULFID 489 504 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT DISULFID 522 533 BY SIMILARITY.
FT DISULFID 527 542 BY SIMILARITY.
FT DISULFID 544 553 BY SIMILARITY.
FT DISULFID 560 571 BY SIMILARITY.
FT DISULFID 565 580 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT DISULFID 598 609 BY SIMILARITY.
FT DISULFID 603 618 BY SIMILARITY.
FT DISULFID 620 629 BY SIMILARITY.
FT DISULFID 636 647 BY SIMILARITY.
FT DISULFID 641 656 BY SIMILARITY.
FT DISULFID 658 667 BY SIMILARITY.
FT DISULFID 674 685 BY SIMILARITY.
FT DISULFID 679 694 BY SIMILARITY.
FT DISULFID 705 723 BY SIMILARITY.
FT DISULFID 712 732 BY SIMILARITY.
FT DISULFID 717 732 BY SIMILARITY.
FT DISULFID 734 743 BY SIMILARITY.
FT DISULFID 750 761 BY SIMILARITY.
FT DISULFID 755 770 BY SIMILARITY.
FT DISULFID 772 781 BY SIMILARITY.
FT DISULFID 788 799 BY SIMILARITY.
FT DISULFID 793 808 BY SIMILARITY.
FT DISULFID 810 819 BY SIMILARITY.
FT DISULFID 826 837 BY SIMILARITY.
FT DISULFID 831 846 BY SIMILARITY.
FT DISULFID 848 857 BY SIMILARITY.
FT DISULFID 864 875 BY SIMILARITY.
FT DISULFID 869 884 BY SIMILARITY.
FT DISULFID 886 895 BY SIMILARITY.
FT DISULFID 902 913 BY SIMILARITY.
FT DISULFID 907 922 BY SIMILARITY.
FT DISULFID 924 933 BY SIMILARITY.
FT DISULFID 936 947 BY SIMILARITY.
FT CAROXYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROXYD 851 851 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARPELIC 477 780 Missing (in isoform IB).
FT CONFLICT 279 279 /FTId=VSP_000451.
SQ SEQUENCE 1064 AA; 112072 MW; 25569CA012ED6D09 CRC64;
Query Match 12.8% Score 120; DB 1; Length 1064;
Beet Local Similarity 28.4% Pred. No. 0.025; Indels 24; Gaps 7;
Matches 40; Conservative 24; Mismatches 53;

OY 34 VSAAGAG---ITGTWYNOLGSTFIYTAGAGALTGTY---ESAVGNAESRYVL---TGR 83
 DB 931 VNEEEVGFCDLEGMNNECHDQVITTKTSGMLGDYMTYNELALYAPTVVVGASANN 990
 OY 84 YDSAPATDGSATGALGWTAMKNNRYNAHSAATTWSGO--YVGAGEARINTOMLTLTS--GTTEA 141
 DB 991 YDF-----PSFGFTV-----RDNGQSTTSWTGCHLCDSEVLVYTTMINNMTSQ 1038
 OY 142 NAMKSTLVGHDTFTKXKPSMA 162
 DB 1039 DIKSNMVGQDKWTRYESQIA 1059

RESULT 11
 ID AVR2_CHICK STANDARD; PRT; 150 AA.
 AC P56732;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Avidin-related protein 2 precursor.
 GN AVR2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxId=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn; TISSUE=Oviduct;
 RX MEDLINE=94170814; PubMed=8125122;
 RA Keiminen R.A., Hallen M.J., Kristo P.A., Laukkanen M.O.,
 RA Tolmela T.A., Helenius M.A., Kulomaa M.S.;
 RT "Molecular cloning and nucleotide sequence of chicken avidin-related
 RT gene 1-5."
 RL Eur. J. Biochem. 220:615-621(1994).
 CC -1- SIMILARITY: BELONGS TO THE AVIDIN/STREPTAVIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; 221554; -; NOT ANNOTATED CDS.
 CC EMBL; 221535; -; NOT ANNOTATED CDS.
 DR PIR; S43202; S43202.
 DR HSP; P02701; IRAY.
 DR Interpro; IPR005469; Avidin.
 DR Interpro; IPR005468; Avidin/str.
 DR Pfam; PF01382; Avidin; 1.
 DR PRINTS; PR00709; AVIDIN.
 DR PROSITE; PS00577; AVIDIN; FALSE_NEG.
 KW Biotin; signal; Multigene family.
 FT SIGNAL 1 24
 FT CHAIN 25 150
 FT BINDING 57 57
 FT CARBOHYD 67 67
 FT CARBOHYD 93 93
 SQ SEQUENCE 150 AA; 16434 MW; 763DIE2BIA3A66D CRC64;

Query Match 12.4%; Score 116.5; DB 1; Length 150;
 Best Local Similarity 32.8%; Pred. No. 0.006;
 Matches 44; Conservative 24; Mismatches 43; Indels 23; Gaps 10;

OY 34 VSAAGAGITGTWYNOLGSTFIYTAGA---DGAITGYESAV---GNASRYVLGRYDS 86
 DB 23 LSAARKSLTGEMNDLGS--IMTIGAVNDNGEDGTITAVADNPENITLSPILGIQHKR 80
 OY 87 APATDGSATGALGWTAMKNNRYNAHSAATTWSGO--YV--GSAERINTOMLTLTSCTTE--AN 142

DB 81 A-----SQPTEFTYHW--NF--SESTSVFVGQCVDSGKGV-LKTKMLQRLAVDDISD 130
 OY 143 AMKSTLVGHDTFTK 156
 DB 131 DWIATRVGNDFTR 144

RESULT 12
 ID SLAP_CAUCR STANDARD; PRT; 1025 AA.
 AC P35828; Q46015; Q9RF12;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE S-layer protein (paracrystalline surface layer protein).
 GN R5AA OR CCI007.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxId=155892;
 RN (1)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21; 646-660 AND 1021-1025.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=93007489; PubMed=1393820;
 RA Gilchrist A., Fisher J.A., Smit J.K.;
 RT "Nucleotide sequence analysis of the gene encoding the Caulobacter
 RT crescentus paracrystalline surface layer protein."
 RL Can. J. Microbiol. 38:193-202(1992).
 RN (2)
 RP REVISIONS TO 376; 636 AND 842-843.
 RA Awram P.;
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=JS3001;
 RA Bingle W.H., Awram P.A., Nornellini J.F., Smit J.K.;
 RT "The secretion signal of C. crescentus S-layer protein is located in
 RT the C-terminal 82 amino acids of the molecule."
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173598; PubMed=11259647;
 RA Nieman W.C., Feildlyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Bisen J., Heideberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolney J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 RN (5)
 RP SEQUENCE OF 1-313 FROM N.A., AND SEQUENCE OF 1-20.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=89008089; PubMed=3049545;
 RA Fisher J.A., Smit J.K., Agabian N.;
 RT "Transcriptional analysis of the major surface array gene of
 RT Caulobacter crescentus."
 RL J. Bacteriol. 170:4706-4713(1988).
 RN (6)
 RP CHARACTERIZATION.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=98292737; PubMed=9620954;
 RA Awram P., Smit J.K.;
 RT "The Caulobacter crescentus paracrystalline S-layer protein is
 RT secreted by an ABC transporter (type I) secretion apparatus."
 CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
 CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. PROBABLY ACTS AS A
 CC PHYSICAL BARRIER TO PARASITES AND LYTIC ENZYMES.
 CC -1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-

```

CC LAYER WITH HEXAGONAL SYMMETRY. SECRETED BY AN ABC TRANSPORTER
CC (TYPE I) SECRETION APPARATUS.
CC -1- MISCELLANEOUS: THE CALCIUM BINDING OF THIS PROTEIN MAY MEDIATE
CC SURFACE ATTACHMENT BY IONIC BRIDGING BETWEEN THE S-LAYER AND A
CC SPECIFIC MEMBRANE-ASSOCIATED OLIGOSACCHARIDE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AF062345; AAC38665.2; -
CC EMBL, AF193063; AAF19365.1; -
CC EMBL, AE005779; AAK22991.1; ALT_INIT.
CC PIR: A48995; A48995.
CC HSSP: P22629; ISWC.
CC TIGR: CCI007; -
CC InterPro: IPR001343; HemIysn_Ca_bind.
CC Pfam: PF00353; hemolysincbind; 3.
CC PRINTS: PR00313; CABNDNGRPT.
CC Cell wall, S-layer; Calcium-binding; Complete proteome.
CC INIT_MER 0 0
CC SEQUENCE 1025 AA; 98001 MW; AD7A326E1363D8AC CRC64;
CC -----
Query Match 10.7%; Score 100.5; DB 1; Length 1025;
Best Local Similarity 28.8%; Pred. No. 0.92;
Matches 55; Conservative 18; Mismatches 85; Indels 33; Gaps 7;
QY 6 VAAIVSLTTSITASASADPSKSKAQV---SAAEAGITGTWYNQIGSTFIYTAGADG 61
DB 238 VQAAAVTALPTGVITSGIETMVTWVGAAITLNTSSGVGLTLALNTGSAQTVIYAGAGQ 357
QY 62 ALTGYESAV-----GNAESRYVLTRGYDSAPLTDGSGTLGWTVMKNNYRNAHSGAT 114
DB 338 NLITATTAQAANNAVVDGGAANYTVASITG-VISGTTVGNASASGTV---SVSYANST 412
QY 115 TWSGOY-VGGAERINTOWLTLSTGTEANAMKSTIVGHD-----TFTVKKPSASI 164
DB 413 TTTGIAIVTGGTA-----VVAQTAGAAVNTTLTDADVITVGNSTTAIVITOTAAAT 465
QY 165 DAAKAGVNNG 175
DB 466 AGATVAGRVNG 476

```

```

RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RT Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN= CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + glyoxylate = S-malate +
CC COA.
CC -1- PATHWAY: Glyoxylate bypass; second step.
CC -1- SUBUNIT: Monomer (By similarity). (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic
CC -1- SIMILARITY: Belongs to the malate synthase family. GlcB subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, Z78020; CAB01465.1; -
CC EMBL, AE007047; AAK46156.1; -
CC PIR: F70722; F70722.
CC PDB: 1N81; 18-DEC-02.
CC PDB: 1N8W; 18-DEC-02.
CC TIGR: MT1885; -
CC Tuberculosis; Rv1837c; -
CC HAMAP: MF_00641; -; 1.
CC InterPro: IPR001465; Malate_synthase.
CC Pfam: PF01274; Malate_synthase; 1.
CC TIGRFAMs: TIGR01345; malate syn G; 1.
CC Transferase; Glyoxylate bypass; Tricarboxylic acid cycle;
CC Complete proteome; 3D-structure.
CC ACT SITE 339 339 CATALYTIC BASE (BY SIMILARITY).
CC ACT SITE 633 633 CATALYTIC ACID (BY SIMILARITY).
CC SEQUENCE 741 AA; 80403 MW; A92F54E0FEB87C64 CRC64;
CC -----
Query Match 10.3%; Score 96.5; DB 1; Length 741;
Best Local Similarity 26.5%; Pred. No. 1.4;
Matches 56; Conservative 28; Mismatches 86; Indels 41; Gaps 11;
QY 9 IAVSLTTSITASASAD---PSKSKAQSAAEAGITGTWYNQIGSTFIYTAGADGALTG 65
DB 100 ITTSGVDAEITTTGPGLVVPLNARFALNAANR-WGSLYDALYGVDPVE-TDGAEGK 157
QY 66 -TYESAVGNAESRYVLTRGYDSAPATDGS-GTALGWT-----VAMKNNYRNAHSGATTW 116
DB 158 PTVKRVGDKVIAVARFELDSVPLSSGSGFDATGFTVQDGLVVALPDKSTGIANGQF 217
QY 117 SGOVVGGAERINTOWLTLSTGTEANAMKSTIVGHD-----GTFVKVXP 159
DB 218 AG-YTGAABE--PTSVLLIHGHILIDPESQVGTIDAGVGVILBSAITTIMDFD 274
QY 160 SAAIDAARKA-----GVNNGNPLDAVOQ 183
DB 275 SVAAVDADAVLVGYRNLGLNKDGLAAAVDK 305

```

```

RESULT 14
ID Y309 MYCGE STANDARD; PRT; 1225 AA.
AC P47551; O49317;
DT 01-FEB-1996 (Rel. 33, Created)

```

DT 01-FEB-1996 (Rel. 33, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Hypochemical lipoprotein MG309 precursor.
 MG309.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmales; Mycoplasma.
 NC NCB1_TaxID=2097;
 RX MEDLINE=96026346; PubMed=565993;
 RA Fraser C.M., Goetz J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium";
 RL Science 270:397-403(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bort K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 sequencing";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 (Potential).
 CC -1- SIMILARITY: BELONGS TO THE MG307 / MG309 / MG338 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U39711; AAC7153.1; -
 DR EMBL: U02200; AAD12488.1; -
 DR PIR: B64234; B64234.
 DR TIGR: MG309; -
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR KW Hypochemical protein; Lipoprotein; Membrane; Signal;
 KM Complete proteome.
 FT SIGNAL 1 27
 FT CHAIN 28 1225
 FT LIPID 28 28
 FT CONFLICT 1185 1185 L -> V (IN REF. 2).
 SO SEQUENCE 1225 AA; 138375 MW; C3E4B5B431B6E8 CRC64;
 Query Match 9.9%; Score 93; DB 1; Length 1225;
 Best Local Similarity 24.6%; Pred. No. 4.6;
 Matches 56; Conservative 22; Mismatches 78; Indels 72; Gaps 9;
 QY 4 IYVAALAVLTTSTIRASASADPS-----KDSKQVSAEAGITGTY-----NQLGS 51
 DB 12 LLLSSIAVSL--GIAVAACAQPNSTRIENLFRPSSAFDDKDGKSNATLYKALEVREG 68
 QY 52 TFIIVTAGDAGALTGTYESAVGNESRYVLTRGYSAPATDGSCTALGWTAMKNNRNH 111
 DB 69 TQVLTIRLAPVLNFEENVDDIKRNLRTFND-----IDNS-----FVNGEQLRN-- 116
 QY 112 SATWSGQYVGAEARINTQWLTSGTTEANAKSTLVG-----HDTFK----- 156
 DB 117 -----QYRGDYVLQTDILDNTGNGQAN--WKLADVNKKIYDDEFINKLFTKNFEYVD 168
 QY 157 -----VKPSASISDAKKGKGVNNGNPLDVAQ 182
 DB 169 KSVGLVSTPLKGLIENQSNMNNIKIQAKFVDKXKRLRINNDAVYAIQ 216

RESULT 15
 ID STFR_ECOLI STANDARD; PRT: 1120 AA.
 AC P76072; P77560;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Side tail fiber protein homolog from lambdaoid prophage Rac.
 GN STFR OR B1372.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NC NCB1_TaxID=562;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sempel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takeuchi K., Takeuchi Y., Wada C., Yamamoto Y., Horiiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map";
 RL DNA Res. 3:363-377(1996).
 CC -1- SIMILARITY: BELONGS TO THE TAIL FIBER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE000234; AAC74454.1; ALT_INIT.
 DR EMBL: D90774; BAA14966.1; -
 DR EMBL: D90775; BAA14975.1; -
 DR PIR: G64887; G64887.
 DR Ecogene; EG13370; stfr.
 DR InterPro; IPR005003; Phage_fiber.
 DR InterPro; IPR005068; Phage_fiber_2.
 DR Pfam; PF03335; Phage_fiber_6.
 DR Pfam; PF03406; Phage_fiber_2; 1.
 DR KW Hypochemical protein; Fiber protein; Repeat; Complete proteome.
 SO SEQUENCE 1120 AA; 113779 MW; 54ZB59D71EB795B4 CRC64;
 Query Match 9.8%; Score 91.5; DB 1; Length 1120;
 Best Local Similarity 24.7%; Pred. No. 5.5;
 Matches 42; Conservative 32; Mismatches 89; Indels 7; Gaps 4;
 QY 10 AVSLTIVSTISASADPSKQVSAEAGITGTYVQLGSGTFTVAGDALTGTYES 69
 DB 219 ATSASTATTKASEATTSARDAASKEAKSSETNA--SSASASSASTAAGNSAKAAKT 276
 QY 70 AVGNESRYVLTRGYSAPATDGSCTALGWTY--AMKNNYRNHASTTWSGQYVGAEARI 128
 DB 277 SETNRSSETTAQO--SASAAAGSKTAAASSASASTSAGQASATTAAGSAESAASSA 334
 QY 129 NTQWLTSGTTE--ANAKSTLVGHDTFTKYVPSAASIDAAKKAGVNNGN 176
 DB 335 STATTKAGEATEQASAAARSASAAKTSETNKASETSAESSKTAASASSA 384

Tue Oct 28 09:01:56 2003

us-09-589-870b-2.rsp

Page 12

Search completed: October 27, 2003, 10:42:31
Job time : 15 secs

...

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 10:40:04 ; Search time 19 Seconds

(without alignments)
926.256 Million cell updates/sec

Title: US-09-589-870B-2

Perfect score: 936
Sequence: 1 MKRIVVAIAVSLTVSITA.....IDAAKKAGVNGNPLDAVQ 183

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	100.0	183	2 A23513	streptavidin precu
2	932	99.6	183	2 S57284	streptavidin v1 pr
3	897	95.8	183	2 S57285	streptavidin v2 pr
4	146.5	15.7	150	2 S42204	avidin-related pro
5	143.5	15.3	152	1 VICH	avidin precursor l
6	121.5	13.0	150	2 S42201	avidin-related pro
7	120.5	12.9	150	2 S42203	avidin-related pro
8	120	12.8	1064	2 A40136	fibropellin 1a - s
9	116.5	12.4	150	2 S42202	avidin-related pro
10	112	12.0	595	2 A48558	flagellin - Escher
11	102	10.9	1310	2 A21380	glycosylase homolo
12	102	10.9	1374	2 A21359	extracellular seri
13	101	10.8	595	2 B48658	flagellin - Escher
14	100.5	10.7	1026	2 A48995	paracrystalline su
15	100.5	10.7	1073	2 C87374	S-layer protein Rs
16	99.5	10.6	1477	2 B43855	high-molecular-wet
17	99.5	10.6	2468	2 A83412	hypothetical prote
18	96.5	10.3	741	2 F70722	probable glcB prot
19	96	10.3	4776	2 E95206	cell wall surface
20	94.5	10.1	507	2 T35677	probable hydrolase
21	94	10.0	570	2 S56132	cellulase (EC 3.2.
22	94	10.0	967	2 S66852	hypothetical prote
23	93	9.9	1225	1 B64234	hypothetical prote
24	92.5	9.9	1018	2 H83135	probable adhesin p
25	92.5	9.9	1461	2 E90696	hypothetical prote
26	92.5	9.9	1461	2 A85547	hypothetical prote
27	92	9.8	1217	2 S52714	sericinB - silkw
28	91.5	9.8	1063	2 D86731	hypothetical prote
29	91.5	9.8	1122	2 G64887	probable tail fibre

30	91.5	9.8	2249	2 A41477	190k surface anti
31	91	9.7	782	2 A10062	conserved hypothet
32	91	9.7	867	2 T45463	membrane glycoprot
33	91	9.7	1035	2 AD3203	autotransporter pr
34	90.5	9.7	2761	2 T21064	hypothetical prote
35	90	9.6	13288	2 T03099	mucin, submaxillar
36	89.5	9.6	971	2 B90835	probable tail fibre
37	89.5	9.6	973	2 C85693	probable membrane
38	89	9.5	255	1 CP8MMU	muramoylpentapepti
39	89	9.5	488	2 A32235	poly(3-hydroxybuty
40	89	9.5	641	2 A12524	hypothetical prote
41	89	9.5	866	2 T45462	membrane glycoprot
42	89	9.5	1999	2 AB2018	hypothetical prote
43	88.5	9.5	131	2 A11540	hypothetical prote
44	88.5	9.5	447	2 C84306	hypothetical prote
45	88.5	9.5	465	2 A47023	S-layer protein -

ALIGNMENTS

```

RESULT 1
A23513
streptavidin precursor - Streptomyces avidinii
C:Species: Streptomyces avidinii
C>Date: 03-Nov-1987 #sequence_revision 03-Nov-1987 #next_change 11-Jan-2002
C/Accession: A23513; S11540
R:Argarana, C.E.; Kuntz, I.D.; Birken, S.; Axel, R.; Cantow, C.R.
Nucleic Acids Res. 14, 1871-1882, 1986
A>Title: Molecular cloning and nucleotide sequence of the streptavidin gene.
A/Reference number: A23513; MUID:86148514; PMID:3951999
A/Accession: A23513
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-183 <ARG>
A/Cross-references: GB:X03591; NID:946740; PIDN:CAA27265.1; PID:946741
R:Gilpin, G.; Bayer, E.A.; Wilchek, M.
Biochem. J. 269, 527-530, 1990
A>Title: Studies on the biotin-binding sites of avidin and streptavidin. Tyrosine resid
A/Reference number: S11540; MUID:90351377; PMID:2386489
A/Accession: S11540
A/Molecule type: protein
A/Residues: 54-66 'X', 68 <GIT>
C:Superfamily: streptavidin
C/Keywords: biotin binding; homotetramer
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-183/Product: streptavidin #status predicted <MAT>
F:67/Binding site: biotin (Tyr) #status predicted
F:78/Binding site: biotin (Tyr) #status predicted

Query Match      100.0%; Score 936; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.88-66;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVVAIAVSLTVSITASADPSKQVSAEAGTGTWYNOLGTFIVAGAD 60
DB 1 MKRIVVAIAVSLTVSITASADPSKQVSAEAGTGTWYNOLGTFIVAGAD 60
QY 61 GALTGTYESAVNGAESRYVLTGRYDAPATDGSFALGTVAMKNNYRAHSAATTWSGQY 120
DB 61 GALTGTYESAVNGAESRYVLTGRYDAPATDGSFALGTVAMKNNYRAHSAATTWSGQY 120
QY 121 VGGAERINTQWLLTSGTTEANAMKSTLVGHDTFTKVPKSAASIDAACKAGVNGNPLDA 180
DB 121 VGGAERINTQWLLTSGTTEANAMKSTLVGHDTFTKVPKSAASIDAACKAGVNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183

RESULT 2
S57284

```

streptavidin v1 precursor - Streptomyces venezuelae
C:Species: Streptomyces venezuelae
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
C:Accession: S57284
R:Bayar, E.A.; Kulik, T.; Adar, R.; Wilchek, M.
Biochim. Biophys. Acta 1263, 60-66, 1995
A:Title: Close similarity among streptavidin-like, biotin-binding proteins from Streptom
A:Reference number: S57284; MUID:95359204; PMID:7632734
C:Accession: S57284
A:Molecule type: DNA
A:Residues: 1-183 <BAV>
A:Experimental source: strain Tue2460
C:Superfamily: streptavidin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-183/Product: streptavidin v1 #status predicted <MAT>

Query Match 99.6%; Score 932; DB 2; Length 183;
Best Local Similarity 99.5%; Pred. No. 3,7e-66;
Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MKRIVAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
DB 1 MKRIVAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
OY 61 GALTGTYSAVGNASRYVLTGRYPADTDSGTLGWTAMKNNYNNAHSATTWSCGY 120
DB 61 GALTGTYSAVGNASRYVLTGRYPADTDSGTLGWTAMKNNYNNAHSATTWSCGY 120
OY 121 VGGAEARINTOWLITSGTTEANAMKSTLVGHDTFTKVPKSAASIDAARKAGVNNGNPLDA 180
DB 121 VGGAEARINTOWLITSGTTEANAMKSTLVGHDTFTKVPKSAASIDAARKAGVNNGNPLDA 180
OY 181 VQO 183
DB 181 VQO 183

RESULT 3
streptavidin v2 precursor - Streptomyces venezuelae
C:Species: Streptomyces venezuelae
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
C:Accession: S57285
R:Bayar, E.A.; Kulik, T.; Adar, R.; Wilchek, M.
Biochim. Biophys. Acta 1263, 60-66, 1995
A:Title: Close similarity among streptavidin-like, biotin-binding proteins from Streptom
A:Reference number: S57284; MUID:95359204; PMID:7632734
C:Accession: S57285
A:Molecule type: DNA
A:Residues: 1-183 <BAV>
A:Experimental source: strain Tue2605
C:Superfamily: streptavidin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-183/Product: streptavidin v2 #status predicted <MAT>

Query Match 95.8%; Score 897; DB 2; Length 183;
Best Local Similarity 95.1%; Pred. No. 2,1e-63;
Matches 174; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 MKRIVAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
DB 1 MKRIVAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
OY 61 GALTGTYSAVGNASRYVLTGRYPADTDSGTLGWTAMKNNYNNAHSATTWSCGY 120
DB 61 GALTGTYSAVGNASRYVLTGRYPADTDSGTLGWTAMKNNYNNAHSATTWSCGY 120
OY 121 VGGAEARINTOWLITSGTTEANAMKSTLVGHDTFTKVPKSAASIDAARKAGVNNGNPLDA 180
DB 121 VGGAEARINTOWLITSGTTEANAMKSTLVGHDTFTKVPKSAASIDAARKAGVNNGNPLDA 180
OY 181 VQO 183
DB 181 VQO 183

DB 181 VQO 183

RESULT 4
avidin-related protein 4/5 precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999
C:Accession: S42204; S42205
R:Keinonen, R.A.; Wallen, M.J.; Kristo, P.A.; Laakkonen, M.O.; Toimela, T.A.; Helenius,
Eur. J. Biochem. 220, 615-621, 1994
A:Title: Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5.
A:Reference number: S42201; MUID:94170814; PMID:8125122
C:Accession: S42204
A:Molecule type: DNA
A:Residues: 1-150 <KEI>
A:Cross-references: EMBL:Z22883; NID:9311811
A:Experimental source: strain White Leghorn; tissue oviduct
A:Genetics: CH1
A:Accession: S42205
A:Molecule type: DNA
A:Residues: 1-150 <KEW>
A:Cross-references: EMBL:Z22882; NID:9311812
A:Experimental source: strain White Leghorn; tissue oviduct
A:Genetics: CH2
A:Genetics: <CH1>
A:Gene: avr4
A:Introns: 27/3; 96/1; 136/2
C:Genetics: <CH2>
A:Gene: avr5
A:Introns: 27/3; 96/1; 136/2
C:Superfamily: avidin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-150/Product: avidin-related protein 4/5 #status predicted <MAT>
F:26-105/Disulfide bonds: #status predicted

Query Match 15.7%; Score 146.5; DB 2; Length 150;
Best Local Similarity 33.8%; Pred. No. 0.00011;
Matches 45; Conservative 22; Mismatches 47; Indels 19; Gaps 8;

OY 34 VSAEAGITGTWYNOLGSTFIYTA-GADGALTGTYSAAV---GNASRYVLTGRYPAD 88
DB 23 LSARKCSLTGKWTNLSIMTIRAVNSRGFTGTLYAVADNPENITLSPLLGIOHRA- 81
OY 89 ATDSSGTLGWTAMKNNYNNAHSATTWSCGYV---GGAEARINTOWLITSGTTEAN-AW 144
DB 82 ----SGTTFGTVM--NF--SESTVFTGOCFIDRNGKEV-LKTWLLRSSVNDISYDW 132
OY 145 KSTLVGHDTFTKV 157
DB 133 KATRVGYNNFTRL 145

RESULT 5

VICH
avidin precursor [validated] - chicken
C:Species: Gallus gallus (chicken)
C>Date: 24-Apr-1994 #sequence_revision 04-Nov-1994 #text_change 15-Sep-2000
C:Accession: A54975; A27518; A92093; A92092; A03160
R:Wallen, M.J.; Laakkonen, M.O.; Kulomaa, M.S.
submitted to Genbank, January 1994
A:Description: Sequence of the chicken egg-white avidin gene.
A:Reference number: A54975
C:Accession: A54975
A:Molecule type: DNA
A:Residues: 1-21, 'S', 23-152 <MAL>
A:Cross-references: GB:L27818; NID:9450255; PID:9451869
A:Experimental source: adult oviduct, strain White Leghorn
A>Note: difference at position 22 may be due to PCR error in gene sequence
R:Gope, M.L.; Keinonen, R.A.; Kristo, P.A.; Connolly, O.M.; Beattie, W.G.; Zarucki-Schu
Nucleic Acids Res. 15, 3595-3606, 1987
A:Title: Molecular cloning of the chicken avidin cDNA.
A:Reference number: A27518; MUID:87203384; PMID:3575102

A:Accession: A27518
A:Molecule type: mRNA
A:Residues: 1-152 <GOP>
A:Cross-references: GB:X05343; NID:963071; PIDN:CAA28954.1; PID:963072
R:Delange, R.J.; Huang, T.S.
J. Biol. Chem. 246, 698-709, 1971
A:Title: Egg white avidin. III. Sequence of the 78-residue middle cyanogen bromide peptide
A:Reference number: A92093; MUID:71107559; PMID:5100763
A:Accession: A92093
A:Molecule type: protein
A:Residues: 25-57, 'T', 59-76, 'E', 78-152 <DEL2>
A:Experimental source: egg white
A:Note: Approximately 50% of the chains have 58-116
R:Huang, T.S.; Delange, R.J.
J. Biol. Chem. 246, 686-697, 1971
A:Title: Egg white avidin. II. Isolation, composition, and amino acid sequences of the
A:Reference number: A92092; MUID:71107557; PMID:5100762
A:Contents: sequences of tryptic peptides
A:Accession: A92092
A:Molecule type: protein
A:Residues: 25-57, 'T', 59-76, 'E', 78-152 <DEL2>
R:Liivah, O.; Susseman, J.
Submitted to the Brookhaven Protein Data Bank, April 1993
A:Reference number: A51448; PDB:2AVI
A:Contents: annotation; X-ray crystallography, 3.0 angstroms, with biotin, residues 27-5
R:Liivah, O.; Bayer, E.A.; Wilchek, M.; Susseman, J.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 5076-5080, 1993
A:Title: Three-dimensional structures of avidin and the avidin-biotin complex.
A:Reference number: A47554; MUID:93281699; PMID:8506353
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Pugliese, L.; Code, A.; Malcovati, M.; Bolognesi, M.
Submitted to the Brookhaven Protein Data Bank, March 1993
A:Reference number: A51622; PDB:1AVD
A:Contents: annotation; X-ray crystallography, 2.7 angstroms, with biotin, residues 27-5
R:Pugliese, L.; Code, A.; Malcovati, M.; Bolognesi, M.
Submitted to the Brookhaven Protein Data Bank, March 1993
A:Reference number: A51623; PDB:1AVE
A:Contents: annotation; X-ray crystallography, 2.8 angstroms, without biotin, residues 2
J. Mol. Biol. 231, 698-710, 1993
A:Title: Three-dimensional structure of the tetragonal crystal form of egg-white avidin
A:Reference number: A54974; MUID:93294833; PMID:8515446
A:Contents: annotation; X-ray crystallography, 2.7 angstroms
C:Genetics:
A:introns: 27/3; 98/1; 138/2
C:Superfamily: avidin
C:Keywords: glycoprotein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-152/Product: avidin #status experimental <MAT>
F:28-107/Disulfide bonds: #status experimental
F:41/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 15.3%; Score 143.5; DB 1; Length 152;
Best Local Similarity 33.3%; Pred. No. 0.0002;
Matches 44; Conservative 19; Mismatches 56; Indels 13; Gaps 7;

QY 34 VSAEAGIGTGTWYNOLGSGFTYTA--GADALTTGTSANGASRVLGRTDSAPATG 92
DB 23 LSAARKSLTGKWNNDGSMNTTGAIVNSRGEFTGTYYAV-TATSNKIKESPLHGTONIN 81
QY 93 SGT--ALGWTAVAKNNYRNASHATWSGOV--GGAEARINQWLTSQTTE-ANAMKS 146
DB 82 KRQPTFGFTVNMK---FSESTTVFTGOCFIRNKEV-LKTMLLRSSVANDIGDMWA 136
QY 147 TLVGHDTFTK 158
DB 137 TRVGINIFRLR 148

RESULT 6
S42201
avidin-related protein 1 precursor - chicken
N:Alternate names: avr1

C:Species: Gallus gallus (chicken)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #ext_change 07-May-1999
C:Accession: S42201
R:Keinaenen, R.A.; Wallen, M.J.; Kristo, P.A.; Laukkanen, M.O.; Toimela, T.A.; Helenius
Eur. J. Biochem. 220, 615-621, 1994
A:Title: Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5.
A:Reference number: S42201; MUID:94170814; PMID:8125122
A:Accession: S42201
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <KEI>
A:Cross-references: EMBL:221611
C:Genetics:
A:introns: 27/3; 96/1; 136/2
C:Superfamily: avidin
C:Keywords: glycoprotein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-150/Product: avidin-related protein 1 #status predicted <MAT>
F:28-105/Disulfide bonds: #status predicted
F:54,67,93/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.0%; Score 121.5; DB 2; Length 150;
Best Local Similarity 32.8%; Pred. No. 0.01;
Matches 44; Conservative 24; Mismatches 43; Indels 23; Gaps 10;

QY 34 VSAEAGIGTGTWYNOLGSGFTYTA--DGALTYEAV---GNASRYLGRYDS 86
DB 23 LSAARKSLTGKWNNDGSMNTTGAIVNSRGEFTGTYYAV--DNPNTIRSPGLGQHKR 80
QY 87 AATDGGGALGWTAVAKNNYRNASHATWSGO-VY--GGAEARINQWLTSQTTE-AN 142
DB 81 A-----CQPTFGFTVNM--NF--SESTSVFVQCFCVDRKSGEV-LTKMLRLAVLDISD 130
QY 143 AMKSLVGHDTFTK 156
DB 131 DKATRVGNNDPTR 144

RESULT 7
S42203
avidin-related protein 3 precursor - chicken
N:Alternate names: avr3 protein
C:Species: Gallus gallus (chicken)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #ext_change 07-May-1999
C:Accession: S42203; S39800
R:Keinaenen, R.A.; Wallen, M.J.; Kristo, P.A.; Laukkanen, M.O.; Toimela, T.A.; Helenius
Eur. J. Biochem. 220, 615-621, 1994
A:Title: Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5.
A:Reference number: S42203; MUID:94170814; PMID:8125122
A:Accession: S42203
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <KEI>
A:Cross-references: EMBL:221612; NID:965432
R:Kunus, T.A.; Wallen, M.J.; Kulomaa, M.S.
Biochim. Biophys. Acta 1216, 441-445, 1993
A:Title: Induction of chicken avidin and related mRNAs after bacterial infection.
A:Reference number: S39799; MUID:94092737; PMID:8268225
A:Accession: S39800
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 71-150 <KUN>
A:Cross-references: EMBL:221536; NID:965429
C:Genetics:
A:introns: 27/3; 96/1; 136/2
C:Superfamily: avidin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-150/Product: avidin-related protein 3 #status predicted <MAT>
F:28-105/Disulfide bonds: #status predicted

Query Match 12.9%; Score 120.5; DB 2; Length 150;
Best Local Similarity 30.3%; Pred. No. 0.012;

Db 81 A-----SQPFPGTIVM--NF--SESTSVFVGQCFVDRSGKEV-LKTKMLQRLAVDDISD 130

Qy 143 AKKSTLVGHDTPTK 156
| : | | : | :
Db 131 DWIATRVGNNDPFR 144

RESULT 10

flagellin - Escherichia coli (strain B1 316-42)
 C:Species: Escherichia coli
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 26-Aug-1999
 C:Accession: A48658
 R:Schoenhals, G.; Whitfield, C.
 J. Bacteriol. 175, 5395-5402, 1993
 A:Title: Comparative analysis of flagellin sequences from Escherichia coli strains possess
 A:Reference number: A48658; MUID:93374833; PMID:8366026
 A:Accession: A48658
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-595 <SCH>
 A:Cross-references: CB:L07389; NID:gl45996; PIDN:AAA23799.1; PID:gl45997
 C:Superfamily: flagellin

Query Match	12.0%	Score 112	DB 2	length 595
Best Local Similarity	26.1%	Pred. No. 0.24		
Matches 52	Conservative 30	Mismatches 87	Indels 30	Gaps 9

[illegible]

RESULT 11

Glycosylase homologs IM02444 [imported] - *Listeria monocytogenes* (strain EGD-e)
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD1380
R:Glaser, F.; Pringle, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Feihl, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Matournam, A.; Mok, C.; Schluepfer, T.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:1169669
A:Accession: AD1380
A:Stratus: preliminary
A:Molecule type: DNA
A:Residues: 1-1310 <GLA>
A:Cross-References: GB:NC_003210; PIDN:CAD00522.1; PID:g16411932; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
C:Gene: IM02444

Query Match	10.9%;	Score 102;	DB 2;	Length 1310;
Best Local Similarity	24.9%;	Pred. No. 3.4;		
Matches	48;	Conservative	23;	Mismatches 52;
				Indels 70;
				Gaps 9;
Qy	12	SLTTSVITASASADPSKSKAQVSAAEAGTGTWYNQSGTFFIVAGADGAL-TGTYESA	70	

```

Db      860 AVTVVKVPAGADAD-----TWITTSASIVOGTYEME 893
QY      71 VGN-----AESRYVLTGRYDAPATDGSGLGVTMAKNN-----YRNHS 112
Db      894 TGSFADTVADKPYAEKTYV--DGYDK---DGAGTITIVANKDSGVNDLTLYKNASSD 947
QY      113 -----ATTMSGQYVGGAERINTOWLTSGTTEANAKSTL---VGHDTFTKV----- 157
Db      948 NQALSTIVNGEYVQOTLLKPTWMSVQSEFLPUSAGKNSISYKVVTDIGDKADOVSIDKV 1007
QY      158 -----KPSASIDA 166
Db      1008 NIGFTPTVAKVEA 1020

```

RESULT 12

extracellular serine proteinase (EC 3.4.21.-) [imported] - *Brucella melitensis* (strain
C/Species: *Brucella melitensis*
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #next_change 01-Feb-2002
C/Accession: AE3259
R/DelVecchio, V.G.; Kapactral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Loa, T.; Ivanova,
I.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes-
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen *Brucella meliten-*
A/Reference number: AD3252; PMID:11756688
A/Accession: AE3259
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-1374 <KUR>
A/Cross-references: GB:AE008917; PIDN:AA151240.1; PID:g1981929; GSPDB:GN00150
A/Experimental source: strain 16M
C/Genetics:
A/Gene: BMEI0058
A/Map position: I
C/Keywords: hydrolase; serine proteinase

Query Match	10.9%	Score 102	DB 2	Length 1374
Best Local Similarity	21.5%	Pred. No. 3.6		
Matches 45	Conservative 29	Mismatches 65	Indels 70	Gaps 7

Qy	20	ASASADSKDSKQAVSAEAGI	---TSTWVNOGS	-----TFYTAAD	----	60
Db	1078	APACADPASSPVACVVTSDNGIMARIGCDYSKSSRLTNMSONITVILIQSGVDKFEY	-----	-----	-----	1137
Qy	61	---GALTGTYESAVGNAESRYVLTGRYDSAPATDGSGLGW	-----	-----	-----	99
Db	1138	EADTGKILIGGINALYSAISR	-----NSPSDDGATTSAMWLGGLTWTYGESGFYVD	-----	-----	1190
Qy	100	---TVAMKNVNRNHS	-----ATTMSQGYVGAEARINTOWLLT	-----	-----S	136
Db	1191	GQAQINNNYNDNDYNSDTAGKGLADDKKATGAVASITETQRFNIGRMVSVTPQAQLMWSKLS	-----	-----	-----	1250
Qy	137	GTTTEANAMKSTLVGHDTFTKYKPSAASID	165	-----	-----	
Db	1251	MDTFNNINMEANVSLNDSDSLGRGVALD	1279	-----	-----	

RESULT 13

```

flagellin - Escherichia coli (strain Su 1242)
C:Species: Escherichia coli
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 26-Aug-1999
C:Accession: B48658
R:Schoenhals, G.; Whitfield, C.
J. Bacteriol. 175, 5395-5402, 1993
A:Title: Comparative analysis of flagellin sequences from Escherichia coli strains possess
A:Reference number: A48658; MUID:93374833; PMID:8366026
A:Accession: B48658
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-595 <SCH>
1;

```

A:Cross-references: GB:L07387; NID:g290436; PIDN:AAA23797.1; PID:g290437
C:Superfamily: flagellin

Query Match 10.8%; Score 101; DB 2; Length 595;
Best Local Similarity 25.6%; Pred. No. 1.8;
Matches 51; Conservative 29; Mismatches 89; Indels 30; Gaps 9;

QY 7 AAIAVSLTTVSTITASASADPSKDSKAQV---SAAEAGITGTWYNQLGST---FIYTAGADGAL 63
DB 343 ASVTWGTTFYFKTGADAG-AATANAGVSTDTASKEVTLNKKVATKOGTAAVAAAGDTSA 401
QY 64 TGTYESAV-----GNAESRYVLTGRYDSAPAT---DGSGLAG-WTVAMK--- 104
DB 402 TTTKSGVQVQYAVPAAGDGTSAKYADNTVSNATATYTDADGEMTTIGSVTTKYSIDA 461
QY 105 NNRYNHSATWVSGQYVG--GAEARINTQWLITS-GTTEANAMKSTLVGHD---TFTKV 157
DB 462 NNGKVTVDSCGSGKYPKVGAEVYVSANGTLTTDATSEGTVTKDPKALDBAIISSIDKF 521
QY 158 KPSAASIDAKKAGVNNGN 176
DB 522 RSLIGAIONRLDSAVTNIN 540

RESULT 14

A48995 paracrystalline surface layer protein RsaA - Caulobacter crescentus

C:Species: Caulobacter crescentus
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A48995
R:Glichrist, A.; Fisher, J.A.; Smit, J.
Can. J. Microbiol. 38, 193-202, 1992
A:Title: Nucleotide sequence analysis of the gene encoding the Caulobacter crescentus pA
A:Reference number: A48995; M0ID:93007489; PMID:1393820
A:Accession: A48995
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-1026 <GIL>
A:Cross-references: GB:AF062345; GB:M22663; GB:M84760; NID:g6064104; PIDN:AAK38665.2; PI
A:Experimental source: CB15A, ATCC 19089
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:116173; NCBIPI:116174)

Query Match 10.7%; Score 100.5; DB 2; Length 1026;
Best Local Similarity 28.8%; Pred. No. 3.4;
Matches 55; Conservative 18; Mismatches 85; Indels 33; Gaps 7;

QY 6 VAAIAVSLTTVSTITASASADPSKDSKAQV---SAAEAGITGTWYNQLGSTFIYTAGADG 61
DB 299 VQAAAVTALPTGVITISIEITMNTVSGAIIITLNTSSGVTGLTALNTNTSGAAQVTAGAGQ 358
QY 62 ALTGYESAV-----GNAESRYVLTGRYDSAPATDGSGLAGWTVAMKNNRYNHSAT 114
DB 359 NLTTATTAQAANNVAVDGGANVTASTG-VTSGTTTVGANSASAGTV---SVSVANSST 413
QY 115 TWVSGQY-VGGAEARINTQWLITSGTTEANAMKSTLVGHD-----TFTKVPASASI 164
DB 414 TTTGALAVTGGTA-----VTVAQTAGNAVNTTLTGADVTVTGNSSTTAVTVQTAAAT 466
QY 165 DAAKAGVNNG 175
DB 467 AGATVAGRVNG 477

RESULT 15

C87374

S-layer protein RsaA [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: C87374

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; M0ID:21173698; PMID:11259647

A:Accession: C87374
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1073 <STO>
A:Cross-references: GB:AE005673; NID:g13422297; PIDN:AAK22991.1; GSPDB:GN00148
C:Genetics:
A:Gene: CCI107

Query Match 10.7%; Score 100.5; DB 2; Length 1073;
Best Local Similarity 28.8%; Pred. No. 3.6;
Matches 55; Conservative 18; Mismatches 85; Indels 33; Gaps 7;

QY 6 VAAIAVSLTTVSTITASASADPSKDSKAQV---SAAEAGITGTWYNQLGSTFIYTAGADG 61
DB 346 VQAAAVTALPTGVITISIEITMNTVSGAIIITLNTSSGVTGLTALNTNTSGAAQVTAGAGQ 405
QY 62 ALTGYESAV-----GNAESRYVLTGRYDSAPATDGSGLAGWTVAMKNNRYNHSAT 114
DB 406 NLTTATTAQAANNVAVDGGANVTASTG-VTSGTTTVGANSASAGTV---SVSVANSST 460
QY 115 TWVSGQY-VGGAEARINTQWLITSGTTEANAMKSTLVGHD-----TFTKVPASASI 164
DB 461 TTTGALAVTGGTA-----VTVAQTAGNAVNTTLTGADVTVTGNSSTTAVTVQTAAAT 513
QY 165 DAAKAGVNNG 175
DB 514 AGATVAGRVNG 524

Search completed: October 27, 2003, 10:40:36
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 10:40:43 : Search time 17 Seconds
(without alignments)
455.464 Million cell updates/sec

Title: US-09-589-870b-2

Perfect score: 936
Sequence: 1 MKRIVVAIAVSLTVSITA.....IDAKKAGVNGNPLDAVQ 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/prodata/2/1aa/6C COMB.pep.*
6: /cgn2_6/prodata/2/1aa/6D COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	936	100.0	183	US-08-831-399-2	Sequence 2, Appl1
2	936	100.0	183	US-09-366-862-2	Sequence 2, Appl1
3	936	100.0	183	US-09-368-772-2	Sequence 2, Appl1
4	936	100.0	183	PCT-US93-05240-14	Sequence 14, Appl1
5	936	100.0	183	5168049-5	Patent No. 5168049
6	928	99.1	183	US-09-382-276-1	Sequence 1, Appl1
7	928	99.1	183	US-09-385-867-1	Sequence 1, Appl1
8	901.5	96.3	186	US-09-382-276-2	Sequence 2, Appl1
9	899.5	96.1	186	US-09-382-276-3	Sequence 3, Appl1
10	846	90.4	435	US-08-491-988-5	Sequence 5, Appl1
11	834	89.1	159	US-08-948-097-17	Sequence 17, Appl1
12	834	89.1	159	US-08-941-100-5	Sequence 2, Appl1
13	828	88.5	159	US-09-381-430-2	Sequence 2, Appl1
14	828	88.5	159	US-08-628-540-1	Sequence 1, Appl1
15	828	88.5	159	US-08-941-100-1	Sequence 1, Appl1
16	745	79.6	415	US-08-491-988-7	Sequence 7, Appl1
17	681.5	72.1	402	US-08-491-988-9	Sequence 9, Appl1
18	675	72.1	128	US-08-211-833-2	Sequence 2, Appl1
19	675	72.1	128	US-08-334-718-2	Sequence 2, Appl1
20	675	72.1	128	US-08-831-399-16	Sequence 16, Appl1
21	675	72.1	128	US-09-366-862-16	Sequence 16, Appl1
22	675	72.1	128	US-09-368-772-16	Sequence 16, Appl1
23	634	67.7	118	US-07-780-717C-7	Sequence 7, Appl1
24	480	51.3	128	US-09-285-867-4	Sequence 4, Appl1
25	143.5	15.3	152	US-08-831-399-4	Sequence 4, Appl1
26	143.5	15.3	152	US-09-366-862-4	Sequence 4, Appl1
27	143.5	15.3	152	US-09-368-772-4	Sequence 4, Appl1

28	110.5	11.8	892	4	US-09-336-447A-5	Sequence 5, Appl1
29	108	11.5	21	4	US-08-566-921-2	Sequence 2, Appl1
30	105.5	11.3	2736	4	US-09-352-991A-30227	Sequence 30227, A
31	103	11.0	1005	4	US-09-206-942-41	Sequence 41, Appl1
32	103	11.0	1011	4	US-09-206-942-39	Sequence 39, Appl1
33	101.5	10.8	1026	2	US-08-614-377A-7	Sequence 7, Appl1
34	101.5	10.8	1026	3	US-09-142-648B-7	Sequence 7, Appl1
35	101	10.8	20	1	US-08-211-833-3	Sequence 3, Appl1
36	101	10.8	20	1	US-08-434-718-3	Sequence 3, Appl1
37	100.5	10.7	1026	4	US-08-194-290-7	Sequence 7, Appl1
38	99.5	10.6	1036	4	US-09-206-942-73	Sequence 73, Appl1
39	99.5	10.6	1477	1	US-08-038-882-4	Sequence 4, Appl1
40	99.5	10.6	1477	1	US-08-302-832-4	Sequence 4, Appl1
41	99.5	10.6	1477	2	US-08-530-198-4	Sequence 4, Appl1
42	99.5	10.6	1477	2	US-08-469-880-4	Sequence 4, Appl1
43	99.5	10.6	1477	2	US-08-728-470-4	Sequence 4, Appl1
44	99.5	10.6	1477	2	US-08-617-697-4	Sequence 4, Appl1
45	99.5	10.6	1477	3	US-08-719-641-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1

US-08-831-399-2 : Sequence 2, Application US/08831399

Patent No. 6312916

GENERAL INFORMATION:

APPLICANT: Kopeckzi, Erhard; Muller, Rainer;

APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans

TITLE OF INVENTION: Recombinant Inactive Core

TITLE OF INVENTION: Streptavidin Mutants

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESS: Felte & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/831,399

FILING DATE: 1-April-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 196 13 053.0

FILING DATE: 1-April-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 196 37 718.8

FILING DATE: 16-September-1996

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 6312916man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: HUBR 1105

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 183 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-831-399-2

Query Match

Best Local Similarity 100.0%; Score 936; DB 4; Length 183;

Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVAAIAVSLTIVSITASADPSKSKAQAASAEAGITGTWYNOLGSTFIYTAGAD 60
DB 1 MKRIVAAIAVSLTIVSITASADPSKSKAQAASAEAGITGTWYNOLGSTFIYTAGAD 60
QY 61 GALTGTYSAVGNAESRYVLTGRYDSAPATDGSCTALGWTVAWKNNYNNHSAATTWSQGY 120
DB 61 GALTGTYSAVGNAESRYVLTGRYDSAPATDGSCTALGWTVAWKNNYNNHSAATTWSQGY 120
QY 121 VGGARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
DB 121 VGGARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183

RESULT 2
US-09-366-862-2
; Sequence 2, Application US/09366862
; Patent No. 6391571
; GENERAL INFORMATION:
; APPLICANT: Kopeczki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/366,862
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,399
; FILING DATE: 1-April-1997
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6391571man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-366-862-2

Query Match 100.0%; Score 936; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.5e-87;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVAAIAVSLTIVSITASADPSKSKAQAASAEAGITGTWYNOLGSTFIYTAGAD 60
DB 1 MKRIVAAIAVSLTIVSITASADPSKSKAQAASAEAGITGTWYNOLGSTFIYTAGAD 60
QY 61 GALTGTYSAVGNAESRYVLTGRYDSAPATDGSCTALGWTVAWKNNYNNHSAATTWSQGY 120

DB 61 GALTGTYSAVGNAESRYVLTGRYDSAPATDGSCTALGWTVAWKNNYNNHSAATTWSQGY 120
QY 121 VGGARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
DB 121 VGGARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183

RESULT 3
US-09-368-772-2
; Sequence 2, Application US/09368772
; Patent No. 6417331
; GENERAL INFORMATION:
; APPLICANT: Kopeczki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/368,772
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,399
; FILING DATE: 1-April-1997
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6417331man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-368-772-2

Query Match 100.0%; Score 936; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.5e-87;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVAAIAVSLTIVSITASADPSKSKAQAASAEAGITGTWYNOLGSTFIYTAGAD 60
DB 1 MKRIVAAIAVSLTIVSITASADPSKSKAQAASAEAGITGTWYNOLGSTFIYTAGAD 60
QY 61 GALTGTYSAVGNAESRYVLTGRYDSAPATDGSCTALGWTVAWKNNYNNHSAATTWSQGY 120
DB 61 GALTGTYSAVGNAESRYVLTGRYDSAPATDGSCTALGWTVAWKNNYNNHSAATTWSQGY 120
QY 121 VGGARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180

Db 121 VCGAERINTQWLLTSGTTEANAMKSTLVGHDTFTVKVPSAASIDAKKAGVNNGNPLDA 180
QY 181 VQQ 183
Db 181 VQQ 183

RESULT 4
PCT-US93-05240-14
Sequence 14, Application PC/TUS9305240
GENERAL INFORMATION:
APPLICANT: NAGARAJAN, VASANTHA
TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN FROM BACILLUS
TITLE OF INVENTION: SUBTILLIS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: DU PONT COMPANY
STREET: BARLEY MILL PLAZA 36
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: USA
ZIP: 19880-0036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05240
FILING DATE: 19930527
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GEIGER, KATHLEEN W
REFERENCE/DOCKET NUMBER: CR 9029
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-2118
TELEFAX: 302-892-7949
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 183 amino acids
TYPE: AMINO ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US93-05240-14

Query Match 100.0%; Score 936; DB 5; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.5e-87;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVVAIAVSLTTVSITASADPSKDSKAQVSAEAGITGTWYNQSGTFTVTAGAD 60
Db 1 MKRIVVAIAVSLTTVSITASADPSKDSKAQVSAEAGITGTWYNQSGTFTVTAGAD 60
QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDSSGTALGWTVMKNNYRNAHSAATTWSGOY 120
Db 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDSSGTALGWTVMKNNYRNAHSAATTWSGOY 120
QY 121 VCGAERINTQWLLTSGTTEANAMKSTLVGHDTFTVKVPSAASIDAKKAGVNNGNPLDA 180
Db 121 VCGAERINTQWLLTSGTTEANAMKSTLVGHDTFTVKVPSAASIDAKKAGVNNGNPLDA 180
QY 181 VQQ 183
Db 181 VQQ 183

RESULT 5
5168049-5
Patent No. 5168049
APPLICANT: MEADE, HARRY M.; GARWIN, JEFFREY L.
TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN-LIKE
POLYPEPTIDES

NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/185,329
FILING DATE: 21-APR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 656,873
FILING DATE: 02-OCT-1984
SEQ ID NO: 5
LENGTH: 183
5168049-5

Query Match 100.0%; Score 936; DB 6; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.5e-87;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVVAIAVSLTTVSITASADPSKDSKAQVSAEAGITGTWYNQSGTFTVTAGAD 60
Db 1 MKRIVVAIAVSLTTVSITASADPSKDSKAQVSAEAGITGTWYNQSGTFTVTAGAD 60
QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDSSGTALGWTVMKNNYRNAHSAATTWSGOY 120
Db 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDSSGTALGWTVMKNNYRNAHSAATTWSGOY 120
QY 121 VCGAERINTQWLLTSGTTEANAMKSTLVGHDTFTVKVPSAASIDAKKAGVNNGNPLDA 180
Db 121 VCGAERINTQWLLTSGTTEANAMKSTLVGHDTFTVKVPSAASIDAKKAGVNNGNPLDA 180
QY 181 VQQ 183
Db 181 VQQ 183

RESULT 6
US-09-382-276-1
Sequence 1, Application US/09382276
Patent No. 6413934
GENERAL INFORMATION:
APPLICANT: Strayton, Patrick S.
APPLICANT: McDevitt, Todd C.
TITLE OF INVENTION: Streptavidin Mutants Having Secondary Functional
FILE REFERENCE: UMS 104
CURRENT APPLICATION NUMBER: US/09/382,276
CURRENT FILING DATE: 1999-08-25
EARLIER APPLICATION NUMBER: 60/097,816
EARLIER FILING DATE: 1998-08-25
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1
LENGTH: 183
TYPE: PRT
ORGANISM: Streptomyces avidinii
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(183)
OTHER INFORMATION: Wild Type Streptavidin
NAME/KEY: PROPEP
LOCATION: (1)..(24)
FEATURE:
NAME/KEY: CHAIN
LOCATION: (25)..(183)
US-09-382-276-1

Query Match 99.1%; Score 928; DB 4; Length 183;
Best Local Similarity 99.5%; Pred. No. 3.6e-86;
Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRIVVAIAVSLTTVSITASADPSKDSKAQVSAEAGITGTWYNQSGTFTVTAGAD 60
Db 1 MKRIVVAIAVSLTTVSITASADPSKDSKAQVSAEAGITGTWYNQSGTFTVTAGAD 60

QY 61 GALTGTYESAVGNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
DB 61 GALTGTYESAVGNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
QY 121 VGGAEARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDA 180
DB 121 VGGAEARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183

RESULT 7
US-09-285-867-1
Sequence 1, Application US/09285867
Patent No. 6492492
GENERAL INFORMATION:
APPLICANT: Stayton, Patrick S.
TITLE OF INVENTION: Circularly Permuted Biotin Binding Proteins
FILE REFERENCE: UMS 103
CURRENT APPLICATION NUMBER: US/09/285,867
CURRENT FILING DATE: 1999-04-02
EARLIER APPLICATION NUMBER: 60/080,560
EARLIER FILING DATE: 1998-04-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 183
TYPE: PRT
ORGANISM: Streptavidin
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(183)
US-09-285-867-1

Query Match 99.1%; Score 928; DB 4; Length 183;
Best Local Similarity 99.5%; Pred. No. 3,6e-86;
Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRIVVAIAVSLTIVSITASASADPSKSKAQSAAEAGITGTWYNOLGSTFTVTAGAD 60
DB 1 MKRIVVAIAVSLTIVSITASASADPSKSKAQSAAEAGITGTWYNOLGSTFTVTAGAD 60
QY 61 GALTGTYESAVGNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
DB 61 GALTGTYESAVGNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
QY 121 VGGAEARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDA 180
DB 121 VGGAEARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183

RESULT 8
US-09-382-276-2
Sequence 2, Application US/09382276
Patent No. 6413934
GENERAL INFORMATION:
APPLICANT: Stayton, Patrick S.
APPLICANT: McDevitt, Todd C.
APPLICANT: Nelson, Kjell J.
TITLE OF INVENTION: Streptavidin Mutants Having Secondary Functional
FILE REFERENCE: UMS 104
CURRENT APPLICATION NUMBER: US/09/382,276
CURRENT FILING DATE: 1999-08-25
EARLIER APPLICATION NUMBER: 60/097,816
EARLIER FILING DATE: 1998-08-25
NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 186
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: FN-SA
FEATURE:
NAME/KEY: PROPEP
LOCATION: (1)..(24)
FEATURE:
NAME/KEY: CHAIN
LOCATION: (25)..(186)
US-09-382-276-2

Query Match 96.3%; Score 901.5; DB 4; Length 186;
Best Local Similarity 96.2%; Pred. No. 1.7e-83;
Matches 179; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MKRIVVAIAVSLTIVSITASASADPSKSKAQSAAEAGITGTWYNOLGSTFTVTAGAD 60
DB 1 MKRIVVAIAVSLTIVSITASASADPSKSKAQSAAEAGITGTWYNOLGSTFTVTAGAD 60
QY 61 GALTGTYESAVGNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWS 117
DB 61 GALTGTYESAVGNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWS 120
QY 118 GOVVGAEARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNP 177
DB 121 GOVVGAEARINTOWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNP 180
QY 178 LDAVQO 183
DB 181 LDAVQO 186

RESULT 9
US-09-382-276-3
Sequence 3, Application US/09382276
Patent No. 6413934
GENERAL INFORMATION:
APPLICANT: Stayton, Patrick S.
APPLICANT: McDevitt, Todd C.
APPLICANT: Nelson, Kjell J.
TITLE OF INVENTION: Streptavidin Mutants Having Secondary Functional
FILE REFERENCE: UMS 104
CURRENT APPLICATION NUMBER: US/09/382,276
CURRENT FILING DATE: 1999-08-25
EARLIER APPLICATION NUMBER: 60/097,816
EARLIER FILING DATE: 1998-08-25
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 186
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: OSTRP-SA
FEATURE:
NAME/KEY: PROPEP
LOCATION: (1)..(24)
FEATURE:
NAME/KEY: CHAIN
LOCATION: (25)..(186)
US-09-382-276-3

Query Match 96.1%; Score 899.5; DB 4; Length 186;
Best Local Similarity 96.2%; Pred. No. 2.8e-83;
Matches 179; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MKRIVVAIAVSLTIVSITASASADPSKSKAQSAAEAGITGTWYNOLGSTFTVTAGAD 60
DB 1 MKRIVVAIAVSLTIVSITASASADPSKSKAQSAAEAGITGTWYNOLGSTFTVTAGAD 60

Db 1 MRKIVAAIAVSLTIVSITASASADPSKQVSAEAGITGTWYNQLSGTFIVTAGD 60
QY 61 GALTGYESAIVGAESRYVLTGRYDAPAAA--TDGSGTALGTVAMKNNYRNAHSATTT 117
Db 61 GALTGYESAIVGAESRYVLTGRYDAPAAA--TDGSGTALGTVAMKNNYRNAHSATTT 120
QY 118 GQYVGAEARINTQWLTSGTTEANAMKSTLVGHDTFTKVPASASIDAKKAGVNNGP 177
Db 121 GQYVGAEARINTQWLTSGTTEANAMKSTLVGHDTFTKVPASASIDAKKAGVNNGP 180
QY 178 LDAVQO 183
Db 181 LDAVQO 186

RESULT 10
US-08-491-988-5
Sequence 5, Application US/08491988
Patent No. 5973116
GENERAL INFORMATION:
APPLICANT: EPIENETOS, AGAMENON A.
APPLICANT: SPOONER, ROBERT A.
APPLICANT: DEONARAIN, MAHENDRA
TITLE OF INVENTION: Compounds for targeting
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: MCAULAY NISSEN GOLDBERG KIEL & HAND, LLP
STREET: 261 MADISON AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/491,988
FILING DATE: 18-DEC-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: GOLDBERG, JUPES E.
REGISTRATION NUMBER: 24,408
TELEPHONE: 212-986-4090
TELEFAX: 212-818-9479
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-491-988-5

Query Match 90.4%; Score 846; DB 2; Length 435;
Best Local Similarity 98.8%; Pred. No. 2.2e-77;
Matches 162; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 20 ASASAPSKDSKAQVSAEAGITGTWYNQLSGTFIVTAGADGALTGTYESAVNAESRYV 79
Db 272 AAAPAPSKDSKAQVSAEAGITGTWYNQLSGTFIVTAGADGALTGTYESAVNAESRYV 331
QY 80 LTRGYSPATDSSGTLGTVAMKNNYRNAHSATTTSGQYVGAEARINTQWLTSGTT 139
Db 332 LTRGYSPATDSSGTLGTVAMKNNYRNAHSATTTSGQYVGAEARINTQWLTSGTT 391
QY 140 EAAMKSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDAVQO 183
Db 392 EAAMKSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDAVQO 435

RESULT 11

US-08-948-097-17
Sequence 17, Application US/08948097C
Patent No. 6103493
GENERAL INFORMATION:
APPLICANT: Skeira, Arne
APPLICANT: Voss, Selma
TITLE OF INVENTION: Streptavidin Mucine
FILE REFERENCE: HUBB 119
CURRENT APPLICATION NUMBER: US/08/948,097C
EARLIER FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 17
SEQ ID NO 17
LENGTH: 159
TYPE: PRP
ORGANISM: Streptomyces avidinii
US-08-948-097-17

Query Match 89.1%; Score 834; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.2e-77;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DSKDSKAQVSAEAGITGTWYNQLSGTFIVTAGADGALTGTYESAVNAESRYVLTGRY 84
Db 1 DSKDSKAQVSAEAGITGTWYNQLSGTFIVTAGADGALTGTYESAVNAESRYVLTGRY 60
QY 85 DSAPATDSSGTLGTVAMKNNYRNAHSATTTSGQYVGAEARINTQWLTSGTTEANAM 144
Db 61 DSAPATDSSGTLGTVAMKNNYRNAHSATTTSGQYVGAEARINTQWLTSGTTEANAM 120
QY 145 KSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDAVQO 183
Db 121 KSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDAVQO 159

RESULT 12
US-08-941-100-5
Sequence 5, Application US/08941100B
Patent No. 6207390
GENERAL INFORMATION:
APPLICANT: Cantor, Charles R.
APPLICANT: Sano, Takeshi
TITLE OF INVENTION: Reduced Affinity Streptavidin
FILE REFERENCE: BU-03165
CURRENT APPLICATION NUMBER: US/08/941,100B
CURRENT FILING DATE: 1997-10-03
PRIOR APPLICATION NUMBER: 08/469,353
PRIOR FILING DATE: 1995-06-06
PRIOR APPLICATION NUMBER: 08/420,010
PRIOR FILING DATE: 1995-04-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 159
TYPE: PRP
ORGANISM: Streptomyces avidinii
US-08-941-100-5

Query Match 89.1%; Score 834; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.2e-77;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DSKDSKAQVSAEAGITGTWYNQLSGTFIVTAGADGALTGTYESAVNAESRYVLTGRY 84
Db 1 DSKDSKAQVSAEAGITGTWYNQLSGTFIVTAGADGALTGTYESAVNAESRYVLTGRY 60
QY 85 DSAPATDSSGTLGTVAMKNNYRNAHSATTTSGQYVGAEARINTQWLTSGTTEANAM 144
Db 61 DSAPATDSSGTLGTVAMKNNYRNAHSATTTSGQYVGAEARINTQWLTSGTTEANAM 120
QY 145 KSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDAVQO 183

Db 121 KSTLVGHDTFTKVPASASIDAAKKAQVNGNPLDAVQ 159

RESULT 13

US-09-381-430-2
Sequence 2, Application US/09381430

Patent No. 6368813

GENERAL INFORMATION:

APPLICANT: Reznik, Gabriel O.

APPLICANT: Sano, Takeshi

APPLICANT: Vajda, Sandor

APPLICANT: Smith, Cassandra

APPLICANT: Cantor, Charles

TITLE OF INVENTION: MULTIFLAVOR STREPTAVIDIN

FILE REFERENCE: 1586-50152

CURRENT APPLICATION NUMBER: US/09/381,430

PRIOR FILING DATE: 2000-03-23

PRIOR APPLICATION NUMBER: PCT/US98/04931

PRIOR FILING DATE: 1998-03-13

PRIOR FILING DATE: 1997-03-14

NUMBER OF SEQ ID NOS: 2

SOFTWARE: Patent in version 3.1

SEQ ID NO 2

LENGTH: 159

TYPE: PRT

ORGANISM: Streptomyces avidinii

US-09-381-430-2

Query Match 89.1%; Score 834; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.2e-77;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 DSKSKAQAQVSAEAGITGTWNOUGSTFIYTAGADGALTGTYESAVNAESRYVLTGRY 84

DB 1 DSKSKAQAQVSAEAGITGTWNOUGSTFIYTAGADGALTGTYESAVNAESRYVLTGRY 60

QY 85 DSAPATDGGTALGWTAVMKNNYRNAHSATTWGGQYVGAEGARINTOWLLTSGTTEANAW 144

DB 61 DSAPATDGGTALGWTAVMKNNYRNAHSATTWGGQYVGAEGARINTOWLLTSGTTEANAW 120

QY 145 KSTLVGHDTFTKVPASASIDAAKKAQVNGNPLDAVQ 183

DB 121 KSTLVGHDTFTKVPASASIDAAKKAQVNGNPLDAVQ 159

RESULT 14

US-08-628-540-1
Sequence 1, Application US/08628540

Patent No. 6022951

GENERAL INFORMATION:

APPLICANT: Sano, Takeshi

APPLICANT: Cantor, Charles R.

APPLICANT: Vajda, Sandor

APPLICANT: Reznik, Gabriel O.

APPLICANT: Smith, Cassandra L.

APPLICANT: Pandori, Mark W.

TITLE OF INVENTION: STREPTAVIDIN MUTANTS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: BAKER & BOTTS, L.L.P.

STREET: 1299 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20004-2400

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/628,540

FILING DATE: 10-APR-1996

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/420,010

FILING DATE: 11-APR-1995

APPLICATION NUMBER: 60/003,687

FILING DATE: 18-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: Remenick, James

REGISTRATION NUMBER: 36,902

REFERENCE/DOCKET NUMBER: 016865-0244

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-639-7700

TELEFAX: 202-639-7890

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 159 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

US-08-628-540-1

Query Match 88.5%; Score 828; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.7e-76;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 PSKSKAQAQVSAEAGITGTWNOUGSTFIYTAGADGALTGTYESAVNAESRYVLTGRYD 85

DB 2 PSKSKAQAQVSAEAGITGTWNOUGSTFIYTAGADGALTGTYESAVNAESRYVLTGRYD 61

QY 86 SAPATDGGTALGWTAVMKNNYRNAHSATTWGGQYVGAEGARINTOWLLTSGTTEANAWK 145

DB 62 SAPATDGGTALGWTAVMKNNYRNAHSATTWGGQYVGAEGARINTOWLLTSGTTEANAWK 121

QY 146 STLVGHDTFTKVPASASIDAAKKAQVNGNPLDAVQ 183

DB 122 STLVGHDTFTKVPASASIDAAKKAQVNGNPLDAVQ 159

RESULT 15

US-08-941-100-1
Sequence 1, Application US/08941100B

Patent No. 6207390

GENERAL INFORMATION:

APPLICANT: Sano, Takeshi

APPLICANT: Cantor, Charles R.

TITLE OF INVENTION: Reduced Affinity Streptavidin

FILE REFERENCE: BU-03165

CURRENT APPLICATION NUMBER: US/08/941,100B

PRIOR FILING DATE: 1997-10-03

PRIOR APPLICATION NUMBER: 08/469,353

PRIOR FILING DATE: 1995-06-06

PRIOR APPLICATION NUMBER: 08/420,010

PRIOR FILING DATE: 1995-04-11

NUMBER OF SEQ ID NOS: 5

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 1

LENGTH: 159

TYPE: PRT

ORGANISM: Streptomyces avidinii

FEATURE:

NAME/KEY: UNSURE

LOCATION: (1)

OTHER INFORMATION: The residue in this position can be any amino

OTHER INFORMATION: acid.

US-08-941-100-1

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 10:42:38 : Search time 28 Seconds
(without alignments)
1094.475 Million cell updates/sec

Title: US-09-589-870b-2

Perfect score: 936
Sequence: 1 MKRIVAAIAVSLTIVSITA.....IDAKKAGVNGNPLDAVQ 183

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	936	100.0	183	US-10-244-821-2	Sequence 2, Appli
2	936	100.0	183	US-10-013-173-2	Sequence 2, Appli
3	936	100.0	183	US-10-150-762-2	Sequence 2, Appli
4	855.5	91.4	412	US-10-244-821-6	Sequence 6, Appli
5	855.5	91.4	412	US-10-013-173-6	Sequence 6, Appli
6	855.5	91.4	412	US-10-150-762-6	Sequence 6, Appli
7	854.5	91.3	431	US-10-244-821-4	Sequence 4, Appli
8	854.5	91.3	431	US-10-013-173-4	Sequence 4, Appli
9	854.5	91.3	431	US-10-150-762-4	Sequence 4, Appli
10	847	90.5	423	US-10-244-821-3	Sequence 8, Appli
11	847	90.5	423	US-10-013-173-3	Sequence 8, Appli
12	847	90.5	423	US-10-150-762-3	Sequence 8, Appli
13	847	90.5	423	US-10-244-821-88	Sequence 8, Appli
14	847	90.5	423	US-10-013-173-88	Sequence 8, Appli
15	847	90.5	423	US-10-150-762-88	Sequence 8, Appli

16	847	90.5	444	15	US-10-150-762-49	Sequence 49, Appli
17	834	89.1	421	10	US-09-938-270B-1	Sequence 1, Appli
18	143.5	15.3	152	10	US-09-957-667-2	Sequence 2, Appli
19	120	12.8	1064	12	US-10-173-461-5	Sequence 5, Appli
20	110.5	11.8	892	11	US-09-952-267-5	Sequence 5, Appli
21	108	11.5	21	15	US-10-289-921-2	Sequence 2, Appli
22	103	11.0	1005	12	US-10-193-764-41	Sequence 41, Appli
23	103	11.0	1011	12	US-10-193-764-39	Sequence 39, Appli
24	103	11.0	4060	12	US-10-140-472-197	Sequence 197, App
25	103	11.0	4060	12	US-10-141-761-197	Sequence 197, App
26	103	11.0	4060	12	US-10-142-885-197	Sequence 197, App
27	103	11.0	4060	12	US-10-158-790-197	Sequence 197, App
28	103	11.0	4060	15	US-10-123-155-197	Sequence 197, App
29	103	11.0	4060	16	US-10-146-731-197	Sequence 197, App
30	102.5	11.0	688	9	US-09-864-761-36047	Sequence 36047, A
31	101.5	10.8	1026	9	US-09-379-931-7	Sequence 7, Appli
32	101.5	10.8	1026	12	US-10-223-597-7	Sequence 7, Appli
33	99.5	10.6	1036	12	US-10-193-764-69	Sequence 69, Appli
34	99.5	10.6	1477	12	US-10-193-764-67	Sequence 67, Appli
35	99.5	10.6	1477	14	US-10-092-880-4	Sequence 4, Appli
36	99.5	10.6	2468	12	US-10-246-330-4	Sequence 4, Appli
37	98	10.5	831	11	US-09-952-267-1	Sequence 1, Appli
38	96.5	10.3	748	11	US-09-272-975-2	Sequence 2, Appli
39	96.5	10.3	753	11	US-09-272-975-58	Sequence 58, Appli
40	96.5	10.3	2933	12	US-10-140-472-345	Sequence 345, App
41	96.5	10.3	2933	12	US-10-141-761-345	Sequence 345, App
42	96.5	10.3	2933	12	US-10-142-885-345	Sequence 345, App
43	96.5	10.3	2933	12	US-10-158-790-345	Sequence 345, App
44	96.5	10.3	2933	15	US-10-123-155-345	Sequence 345, App
45	96.5	10.3	2933	16	US-10-146-731-345	Sequence 345, App

ALIGNMENTS

RESULT 1	US-10-244-821-2	Application US/10244821
Sequence 2, Appli	US20030143233A1	
GENERAL INFORMATION:		
APPLICANT:	Goshorn, Stephen Charles	
APPLICANT:	Graves, Scott Stoll	
APPLICANT:	Schultz, Joanne Elaine	
APPLICANT:	Lin, Yukang	
APPLICANT:	Sanderson, James Allen	
APPLICANT:	Reno, John M.	
APPLICANT:	Dearysne, Erica A.	
TITLE OF INVENTION:	STREPTAVIDIN EXPRESSED GENE FUSIONS AND	
FILE REFERENCE:	690022.547C3	
CURRENT APPLICATION NUMBER:	US/10/244,821	
CURRENT FILING DATE:	2002-09-16	
NUMBER OF SEQ ID NOS:	92	
SOFTWARE:	FastSeq for Windows Version 4.0	
SEQ ID NO 2		
LENGTH:	183	
TYPE:	PRT	
ORGANISM:	Streptomyces avidinii	
US-10-244-821-2		
Query Match	100.0%	Score 936; DB 12; Length 183;
Best Local Similarity	100.0%	Pred. No. 1.7e-85;
Matches	183;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MKRIVAAIAVSLTIVSITASADPSKQSAEAGITGTWYNQIGSTFIYTAGD 60
DB	1	MKRIVAAIAVSLTIVSITASADPSKQSAEAGITGTWYNQIGSTFIYTAGD 60
QY	61	GALTGTYSAGVNASRVLGTGRYDAPATDGSCTALGTVWAKNNYRAHSATTWGCY 120
DB	61	GALTGTYSAGVNASRVLGTGRYDAPATDGSCTALGTVWAKNNYRAHSATTWGCY 120
QY	121	VGAARINTQWLLTSGTTEANAMKSLVGHDTFTKVPASASIDAKKAGVNGNPLDA 180

DB 121 VGGAEARINTQWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
|||
QY 181 VQO 183
|||
DB 181 VQO 183

RESULT 2

US-10-013-173-2
; Sequence 2, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptomyces avidinii
US-10-013-173-2

Query Match 100.0%; Score 936; DB 15; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.7e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVAAIAVSLTIVTSITASASADPSKSKAQAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
|||
DB 1 MKRIVAAIAVSLTIVTSITASASADPSKSKAQAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
|||
QY 61 GALTGYESAVGNAESRYVLTGRYDSAPATDGGTALGWTAMKNNYRNAHSATTWSGOY 120
|||
DB 61 GALTGYESAVGNAESRYVLTGRYDSAPATDGGTALGWTAMKNNYRNAHSATTWSGOY 120
|||
QY 121 VGGAEARINTQWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
|||
DB 121 VGGAEARINTQWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
|||
QY 181 VQO 183
|||
DB 181 VQO 183

RESULT 3

US-10-150-762-2
; Sequence 2, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearsteyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2

; LENGTH: 183
; TYPE: PRT
; ORGANISM: Streptomyces avidinii
US-10-150-762-2

Query Match 100.0%; Score 936; DB 15; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.7e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIVAAIAVSLTIVTSITASASADPSKSKAQAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
|||
DB 1 MKRIVAAIAVSLTIVTSITASASADPSKSKAQAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
|||
QY 61 GALTGYESAVGNAESRYVLTGRYDSAPATDGGTALGWTAMKNNYRNAHSATTWSGOY 120
|||
DB 61 GALTGYESAVGNAESRYVLTGRYDSAPATDGGTALGWTAMKNNYRNAHSATTWSGOY 120
|||
QY 121 VGGAEARINTQWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
|||
DB 121 VGGAEARINTQWLLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDA 180
|||
QY 181 VQO 183
|||
DB 181 VQO 183

RESULT 4

US-10-244-821-6
; Sequence 6, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearsteyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain
; OTHER INFORMATION: antibody-genomic streptavidin fusion
US-10-244-821-6

Query Match 91.4%; Score 855.5; DB 12; Length 412;
Best Local Similarity 96.5%; Pred. No. 5.1e-77;
Matches 165; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 14 TTVSITA-SASADPSKSKAQAQVSAEAGITGTWYNOLGSTFIYTAGADGALTGYESAVG 72
|||
DB 242 TTVTVSSGSGSADPSKSKAQAQVSAEAGITGTWYNOLGSTFIYTAGADGALTGYESAVG 301
|||
QY 73 NAESRYVLTGRYDSAPATDGGTALGWTAMKNNYRNAHSATTWSGOYVGGAEARINTQW 132
|||
DB 302 NAESRYVLTGRYDSAPATDGGTALGWTAMKNNYRNAHSATTWSGOYVGGAEARINTQW 361
|||
QY 133 LITSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDAVQO 183
|||
DB 362 LITSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKKAGVNNGNPLDAVQO 412
|||

RESULT 5
US-10-013-173-6


```

; Sequence 6, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013.173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 412
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain
; US-10-013-173-6

Query Match          91.4%; Score 855.5; DB 15; Length 412;
Best Local Similarity 96.5%; Pred. No. 5.1e-77;
Matches 165; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 14 TTVSITA-SASADPSKDSKAQVSAEAGITGTWYNQLGSTFIYTAGADGALTGTYESAVG 72
    |||:::|
DB 242 TTVTVSSGSGADPSKDSKAQVSAEAGITGTWYNQLGSTFIYTAGADGALTGTYESAVG 301
    |||:::|

QY 73 NAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSCGYVGGAEARINTQW 132
    |||:::|
DB 302 NAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSCGYVGGAEARINTQW 361
    |||:::|

QY 133 LLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDAVOQ 183
    |||:::|
DB 362 LLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDAVOQ 412
    |||:::|

RESULT 6
US-10-150-762-6
; Sequence 6, Application US/10150762.
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearstlyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150.762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 412
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain
; US-10-150-762-6

Query Match          91.4%; Score 855.5; DB 15; Length 412;
Best Local Similarity 96.5%; Pred. No. 5.1e-77;
Matches 165; Conservative 4; Mismatches 1; Indels 1; Gaps 1;
```

```

QY 14 TTVSITA-SASADPSKDSKAQVSAEAGITGTWYNQLGSTFIYTAGADGALTGTYESAVG 72
    |||:::|
DB 242 TTVTVSSGSGADPSKDSKAQVSAEAGITGTWYNQLGSTFIYTAGADGALTGTYESAVG 301
    |||:::|

QY 73 NAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSCGYVGGAEARINTQW 132
    |||:::|
DB 302 NAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSCGYVGGAEARINTQW 361
    |||:::|

QY 133 LLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDAVOQ 183
    |||:::|
DB 362 LLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDAVOQ 412
    |||:::|

RESULT 7
US-10-244-821-4
; Sequence 4, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstlyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244.821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence for hUNR-LU-10
; US-10-244-821-4

Query Match          91.3%; Score 854.5; DB 12; Length 431;
Best Local Similarity 96.5%; Pred. No. 6.8e-77;
Matches 166; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 12 SLTVSITASASADPSKDSKAQVSAEAGITGTWYNQLGSTFIYTAGADGALTGTYESAV 71
    |||:::|
DB 261 TLTVTS-SGSGSADPSKDSKAQVSAEAGITGTWYNQLGSTFIYTAGADGALTGTYESAV 319
    |||:::|

QY 72 GNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSCGYVGGAEARINTQ 131
    |||:::|
DB 330 GNAESRYVLGRYDAPATDGSCTALGWTVMKNNYRNAHSATTWSCGYVGGAEARINTQ 379
    |||:::|

QY 132 WLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDAVOQ 183
    |||:::|
DB 380 WLTSGTTEANAMKSTLVGHDTFTKVPKPSAASIDAAKAGVNNGNPLDAVOQ 431
    |||:::|

RESULT 8
US-10-013-173-4
; Sequence 4, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
```

Query Match	Score	DB	Length
90.5%;	847;	15;	423;

```

; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ. ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 438
; TYPE: PR1
; ORGANISM: Mus musculus
US-10-244-821-88

Query Match 90.5%; Score 847; DB 12; Length 438;
Best Local Similarity 98.2%; Pred. No. 3; 9e-76;
Matches 162; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 19 TASASADPSKDSYAQVSAAEAGITGWNYNOLGSTFIVTGAADALGTYESAVGNAESRY 78
Db 274 SGGSGADPSKDSYAQVSAAEAGITGWNYNOLGSTFIVTGAADALGTYESAVGNAESRY 333
QY 79 VLTGRYDAPATDGSSTALGWTAVMKNNYRNNAHSATTWSGQYVYGAEARINTOMLLTSGT 138
Db 334 VLTGRYDAPATDGSSTALGWTAVMKNNYRNNAHSATTWSGQYVYGAEARINTOMLLTSGT 393
QY 139 TEANAKSTLVGHDTFTKVPSPASASIDAAKKAGVNNGNPLDAVOO 183
Db 394 TEANAKSTLVGHDTFTKVPSPASASIDAAKKAGVNNGNPLDAVOO 438

RESULT 14
US-10-244-821-49
; Sequence 49, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goeshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ. ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 444
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence for the Cc49 single
; OTHER INFORMATION: chain antibody-genomic streptavidin fusion
; OTHER INFORMATION: sequence
US-10-244-821-49

Query Match 90.5%; Score 847; DB 12; Length 444;
Best Local Similarity 98.2%; Pred. No. 4e-76;
Matches 162; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 19 TASASADPSKDSYAQVSAAEAGITGWNYNOLGSTFIVTGAADALGTYESAVGNAESRY 78
Db 280 SGGSGADPSKDSYAQVSAAEAGITGWNYNOLGSTFIVTGAADALGTYESAVGNAESRY 339
QY 79 VLTGRYDAPATDGSSTALGWTAVMKNNYRNNAHSATTWSGQYVYGAEARINTOMLLTSGT 138
Db 340 VLTGRYDAPATDGSSTALGWTAVMKNNYRNNAHSATTWSGQYVYGAEARINTOMLLTSGT 399
QY 139 TEANAKSTLVGHDTFTKVPSPASASIDAAKKAGVNNGNPLDAVOO 183
Db 400 TEANAKSTLVGHDTFTKVPSPASASIDAAKKAGVNNGNPLDAVOO 444

```

```
RESULT 15
US-10-013-173-49
; Sequence 49, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence for the CC49 single
; OTHER INFORMATION: chain antibody-genomic streptavidin fusion
; OTHER INFORMATION: sequence
US-10-013-173-49

Query Match          90.5%; Score 847; DB 15; Length 444;
Best Local Similarity 98.2%; Pred. No. 4e-76;
Matches 162; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 19 TASASADPSKDSKAQVSAEAGITGTWYNQSGSTFIYTAGADGALTGTYESAVGNASRY 78
Db 280 SGSSADPSKDSKAQVSAEAGITGTWYNQSGSTFIYTAGADGALTGTYESAVGNASRY 339
QY 79 VLTGRYSAPATDGGSTALGWTVAWKNNRNNAHSATTWSGGQYVGAEARINTQMLTSGT 138
Db 340 VLTGRYSAPATDGGSTALGWTVAWKNNRNNAHSATTWSGGQYVGAEARINTQMLTSGT 399
QY 139 TEANAMKSTLVGHDTFTKVKPSAASIDAKKAGVNNGNPLDAVQO 183
Db 400 TEANAMKSTLVGHDTFTKVKPSAASIDAKKAGVNNGNPLDAVQO 444
```

Search completed: October 27, 2003, 10:51:47
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 10:40:04 : Search time 42 Seconds
(without alignments)
691.594 Million cell updates/sec

Title: US-09-589-870B-2

Sequence: 1 MKRIIVAAIAVSLTYSIFA.....IDAKKAGVNGNPLDAVQ 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_19jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	936	100.0	183	7 AAP60625	Sequence of a stre
2	936	100.0	183	10 AAP93530	Streptavidin prote
3	936	100.0	183	14 AAR44491	Streptavidin gene.
4	936	100.0	183	18 AAW29306	Wild-type streptav
5	936	100.0	183	19 AAW59216	S. avidinii strept
6	936	100.0	183	20 AAY17668	Streptococcus stre
7	936	100.0	183	21 AAY44701	Streptavidin prote
8	936	100.0	183	22 AAB30692	Amino acid sequenc
9	936	100.0	186	21 AAY80512	Streptomyces avidi

10	931	99.5	184	21 AAY80513	Streptomyces avidi
11	928	99.1	182	8 AAP70492	Streptavidin seque
12	928	99.1	183	21 AAY84020	Amino acid sequenc
13	919	98.2	183	19 AAW59217	S. avidinii strept
14	917	98.0	183	19 AAW59218	Amino acid sequenc
15	901.5	96.3	186	21 AAY84021	Amino acid sequenc
16	899.5	96.1	186	21 AAY84022	Amino acid sequenc
17	855.5	91.4	412	22 AAB30694	A fusion of anti-C
18	854.5	91.3	431	22 AAB30693	A fusion of anti-C
19	847	90.5	423	22 AAB30695	A fusion of anti-C
20	846	90.4	435	15 AAB56483	ScFv PRAS108 and p
21	838	89.5	161	21 AAY80514	Streptomyces avidi
22	835	89.2	162	21 AAY80515	Streptomyces avidi
23	835	89.2	163	9 AAP80160	Biosynthetic prote
24	834	89.1	159	22 AAB35270	Streptomyces avidi
25	834	89.1	160	10 AAP93531	Mature streptavidin
26	834	89.1	421	23 AAM51157	Streptavidin-carbo
27	829	88.6	159	18 AAW29310	Streptavidin prote
28	829	88.6	159	18 AAW29311	Streptavidin prote
29	828	88.5	159	18 AAW29309	Streptavidin prote
30	828	88.5	159	23 AAG78492	Amino acid sequenc
31	824	88.0	159	18 AAW29314	Streptavidin prote
32	823	87.9	159	18 AAW29313	Streptavidin prote
33	823	87.8	159	18 AAW29316	Streptavidin prote
34	821	87.7	159	18 AAW29315	Streptavidin prote
35	820	87.6	159	18 AAW29312	Streptavidin prote
36	820	87.6	159	22 AAB35272	Streptomyces avidi
37	818	87.4	159	18 AAW29319	Streptavidin prote
38	816	87.2	159	18 AAW29318	Streptavidin prote
39	815	87.1	159	18 AAW29317	Streptavidin prote
40	813	86.9	159	23 AAG78495	Mutated version 1
41	810	86.5	159	18 AAW29320	Streptavidin prote
42	790	84.4	159	23 AAG78497	Mutated version 3
43	789	84.3	159	23 AAG78496	Mutated version 2
44	772	82.5	626	24 AAO19549	Biocin-combining s
45	745	79.6	415	15 AAR56484	ScFv PRAS109 and p

ALIGNMENTS

RESULT 1

ID AAP60625 standard; Protein; 183 AA.

AC AAP60625:

DT 25-MAR-2003 (updated)

DT 13-AUG-1991 (first entry)

DE Sequence of a streptavidin-like polypeptide encoded by SA307.

KW Antibiotic; biotin binding affinity; fusion protein.

XX Streptomyces.

OS W08602077-A.

PN 10-APR-1986.

PD 01-OCT-1985; 85WO-0001901.

PF 02-OCT-1984; 84US-0656873.

PR (MEAD/) MEAD H M.

PA (BIOT) BIOGEN NV.

XX Meade HM, Garwin JL, Biogen NV;

PI WPI; 1986-106643/16.

DR N-PSDB; AAN60626.

XX DNA sequences and hybrid DNA sequences - encoding

PT streptavidin-like polypeptide, also joined to another protein,
 PT e.g. tissue plasminogen activator
 XX
 XX Disclosure: Fig. 2; 54pp; English.
 CC The inventors claim the DNA sequence in SA307 which codes for a
 CC streptavidin-like polypeptide (see AAN60626), and the polypeptide
 CC encoded by it (AAP60625). They also claim hybrid SQs comprising AAN60626
 CC and a second sequence coding for another protein, polypeptide,
 CC peptide or AA (pref. tissue plasminogen activator (TPA)).
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX Sequence 183 AA;
 SQ
 Query Match 100.0%; Score 936; DB 7; Length 183;
 Best Local Similarity 100.0%; Pred. No. 9.5e-76;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKRIVVAIAVSLTTSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFTVTAGAD 60
 DB 1 MKRIVVAIAVSLTTSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFTVTAGAD 60
 QY 61 GALTGTYSAGNABSRVLTGRYDSAPATDGGTALGWTWAKNNYRNAHSATTWSGOY 120
 DB 61 GALTGTYSAGNABSRVLTGRYDSAPATDGGTALGWTWAKNNYRNAHSATTWSGOY 120
 QY 121 VGGAERINTQWLTSGTTEANAMKSTLVGHDTFTKVPASAASIDAAKAGVNNGNPLDA 180
 DB 121 VGGAERINTQWLTSGTTEANAMKSTLVGHDTFTKVPASAASIDAAKAGVNNGNPLDA 180
 QY 181 VQQ 183
 DB 181 VQQ 183
 RESULT 2
 AAP93530
 ID AAP93530 standard; protein; 183 AA.
 XX
 AC AAP93530;
 XX
 DT 04-JUN-1990 (first entry)
 XX
 DE Streptavidin protein.
 XX
 KM Streptavidin; Streptomyces avidinii; biotin.
 XX
 OS Streptomyces avidinii.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein /note="Leader sequence"
 FT 25..159
 FT /note="this sequence was as the basis for the design
 FT for the synthetic gene of the present invention."
 FT
 FT
 PN WO8903422-A.
 PN
 PD 20-APR-1989.
 PD
 XX 07-OCT-1988; 88WO-GB00831.
 XX
 XX 08-OCT-1987; 87GB-0023661.
 XX
 XX (BRBI-) BRIT BIO-TECHN LTD.
 XX
 XX Edwards RM;
 XX
 XX WPI; 1989-130040/17.
 XX
 DR DNA sequence encoding streptavidin and vector -
 PT comprising hybrid gene encoding fusion protein with
 PT biotin-binding activity

XX
 PS Fig 1; page 1/5; 22pp; English.
 XX
 CC Streptavidin is a 60KD protein isolated from Streptomyces avidinii that
 CC binds extremely tightly to the vitamin biotin. It is composed of four
 CC identical subunits of 15kd and binds 4 mole of biotin per mole of
 CC protein. It is structurally related to the protein avidin. It can be
 CC readily conjugated to a range of other proteins. In order to facilitate
 CC the incorporation of streptavidin into expression vectors and the
 CC production of novel chimeric proteins containing streptavidin
 CC functionality, an improved novel synthetic gene for streptavidin has
 CC been constructed (AAN90755) based on the amino acid sequence of mature
 CC streptavidin.
 CC
 XX Sequence 183 AA;
 SQ
 Query Match 100.0%; Score 936; DB 10; Length 183;
 Best Local Similarity 100.0%; Pred. No. 9.5e-76;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKRIVVAIAVSLTTSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFTVTAGAD 60
 DB 1 MKRIVVAIAVSLTTSITASADPSKDSKAQVSAEAGITGTWYNOLGSTFTVTAGAD 60
 QY 61 GALTGTYSAGNABSRVLTGRYDSAPATDGGTALGWTWAKNNYRNAHSATTWSGOY 120
 DB 61 GALTGTYSAGNABSRVLTGRYDSAPATDGGTALGWTWAKNNYRNAHSATTWSGOY 120
 QY 121 VGGAERINTQWLTSGTTEANAMKSTLVGHDTFTKVPASAASIDAAKAGVNNGNPLDA 180
 DB 121 VGGAERINTQWLTSGTTEANAMKSTLVGHDTFTKVPASAASIDAAKAGVNNGNPLDA 180
 QY 181 VQQ 183
 DB 181 VQQ 183
 RESULT 3
 AAR44491
 ID AAR44491 standard; Protein; 183 AA.
 XX
 AC AAR44491;
 XX
 DT 25-MAR-2003 (updated)
 XX
 DT 27-JUN-1994 (first entry)
 XX
 DE Streptavidin gene.
 XX
 KM Streptavidin; protein secretion; Bacillus subtilis.
 XX
 OS Streptomyces avidinii.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..24
 FT Misc-difference //label= signal_peptide
 FT Protein 25..183
 FT Peptide /label= streptavidin
 FT 37..183
 FT /note= "expressed by transformed B. subtilis"
 FT
 FT
 PN WO9324631-A1.
 PN
 PD 09-DEC-1993.
 PD
 XX 27-MAY-1993; 93WO-US05240.
 XX
 XX 29-MAY-1992; 92US-0891524.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Nagarajan V;
 XX

DR WPI; 1993-405822/50.
 DR P-PSDB; AAO53412.
 XX
 PT Streptavidin prodn. from Bacillus subtilis - using signal protein
 PT from bacterial exo-protein and expression element from Gram
 PT positive bacterial protein.
 XX
 XX Disclosure; Fig 1b; 54pp; English.
 CC Tetrameric biologically active streptavidin is produced by secretion
 CC from Bacillus subtilis transformed with a plasmid encoding the
 CC sequence.
 CC (updated on 25-MAR-2003 to correct PN field.)
 CC
 XX
 SO Sequence 183 AA;
 Query Match 100.0%; Score 936; DB 14; Length 183;
 Best Local Similarity 100.0%; Pred. No. 9.5e-76;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 MKRIVVAALVSLTTSITASADPSKSKAQSAAEAGITGTWYNQLGSTFIIVAGAD 60
 QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
 DB 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
 QY 121 VGGAERINTOMLLTSGTTEANAMKSTLVGHDTFTKVKPSAASIDAKKAGVNNGNPLDA 180
 DB 121 VGGAERINTOMLLTSGTTEANAMKSTLVGHDTFTKVKPSAASIDAKKAGVNNGNPLDA 180
 QY 181 VQQ 183
 DB 181 VQQ 183
 RESULT 4
 AAW29306
 ID AAW29306 standard; Protein; 183 AA.
 AC AAW29306;
 XX
 DT 27-APR-1998 (first entry)
 XX
 DE Wild-type streptavidin protein.
 XX
 KW Streptavidin; biotin; anti-interference reagent; detection; mutein;
 KM avidin; non-specific binding.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal
 FT Protein 25..183
 XX
 XX DE19637718-A1.
 XX
 XX 02-OCT-1997.
 XX
 XX 16-SEP-1996; 96DE-1037718.
 XX
 XX 01-APR-1996; 96DE-1013053.
 XX
 XX (BOEP) BOEHRINGER MANNHEIM GMBH.
 XX
 XX Brandeteeter H, Deger A, Engh R, Kopetzki E, Mueller R;
 PI Schmitt U;
 XX
 XX WPI; 1997-482043/45.
 DR N-PSDB; AAT73193.
 XX

PT Streptavidin and avidin muteins with reduced binding affinity for
 PT biotin - useful for reducing interference from nonspecific binding
 PT in assays
 XX
 XX Disclosure; Page 17-18; 26pp; German.
 XX
 CC This sequence represents a streptavidin which is used in a novel method
 CC of reducing interference from non-specific binding in assays. Muteins
 CC constructed from a core streptavidin or avidin sequence are selected that
 CC differ from the native polypeptide by at least one amino acid and have a
 CC binding affinity for biotin of less than 1010 1/mole. The biotin-bindable
 CC polypeptide may be present as a polymeric conjugate, e.g. with another
 CC polypeptide or protein, especially bovine serum albumin. These muteins
 CC are used as anti-interference reagents for reducing and/or avoiding
 CC nonspecific interactions in a process for detecting an analyte. In
 CC particular, they are used in assays where the streptavidin/avidin-biotin
 CC specific binding pair is involved for qualitative and/or quantitative
 CC determination of an analyte in a test sample, e.g. a heterogeneous
 CC immunoassay or a hybridisation assay. Despite having a lower binding
 CC affinity for biotin, the muteins have high immunological cross-reactivity
 CC with native streptavidin and avidin.
 CC
 XX
 SO Sequence 183 AA;
 Query Match 100.0%; Score 936; DB 18; Length 183;
 Best Local Similarity 100.0%; Pred. No. 9.5e-76;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1 MKRIVVAALVSLTTSITASADPSKSKAQSAAEAGITGTWYNQLGSTFIIVAGAD 60
 QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
 DB 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSCTALGWTVMKNNYRNAHSATTWSGOY 120
 QY 121 VGGAERINTOMLLTSGTTEANAMKSTLVGHDTFTKVKPSAASIDAKKAGVNNGNPLDA 180
 DB 121 VGGAERINTOMLLTSGTTEANAMKSTLVGHDTFTKVKPSAASIDAKKAGVNNGNPLDA 180
 QY 181 VQQ 183
 DB 181 VQQ 183
 RESULT 5
 AAW59216
 ID AAW59216 standard; Protein; 183 AA.
 AC AAW59216;
 XX
 DT 27-AUG-1998 (first entry)
 XX
 DE S. avidinii streptavidin protein.
 XX
 KW Streptavidin; ligand; binding affinity; mutant; isolation;
 KM purification; recover; immobilise.
 XX
 XX Streptomyces avidinii.
 XX
 XX EP835934-A2.
 XX
 XX 15-APR-1998.
 XX
 XX 09-OCT-1997; 97EP-0117504.
 XX
 XX 10-OCT-1996; 96DE-1041876.
 XX
 XX (BIOA-) INST BIOANALYTIK GMBH GOETTINGEN.
 PA
 XX Skerra A, Voess S;
 PI
 XX WPI; 1998-218868/20.
 XX

DR N-PSDB; AAV34714.

Streptavidin mutants with higher binding affinity for peptide ligands - have mutation in amino acid region 44-53, used to isolate, purify or determine fusion proteins including these ligands

PS Disclosure; Page -; 21pp; German.

CC This sequence encodes a wild-type streptavidin protein isolated from
CC Streptomyces avidinii. This sequence is used to produce mutants which
CC are used in a method to assay the binding affinity of streptavidin
CC mutants. These mutants have a mutation within the amino acid (aa) region
CC 44-53 of the wild-type protein show a higher binding affinity than the
CC wild-type for peptide ligands that include the sequence of formula
CC $\text{Trip-X-His-Pro-Gln-Phe-Y-Z}$ where X = any aa; Y and Z are both Gly,
CC or Y = Glu and Z = Arg or Lys. Recombinant streptavidin mutants can
CC be used to isolate, purify and determine proteins or to
CC determine/recover substances that contain streptavidin-binding groups.
CC Such compounds may also be used to immobilise fusions on microtitre
CC plates, microbeads or sensor chips.
CC NOTE: This sequence does not appear in the specification but is used to
CC make the mutant streptavidin proteins represented in AAV3217 and
CC AAV59218.

CC Sequence 183 AA;

Query Match 100.0%; Score 936; DB 19; Length 183;
Best Local Similarity 100.0%; Pred. No. 9.5e-76;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIIVAAIAVSLTIVSITASASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
DB 1 MKRIIVAAIAVSLTIVSITASASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
QY 61 GALTGTYESAVGNASERIVLTGRYDSAPATDGSCTALGWTVMKNNYNRAHSATTWSGOY 120
DB 61 GALTGTYESAVGNASERIVLTGRYDSAPATDGSCTALGWTVMKNNYNRAHSATTWSGOY 120
QY 121 VGGAERINTOMLITSGTTEANAMKSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDA 180
DB 121 VGGAERINTOMLITSGTTEANAMKSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183

RESULT 6

AAV17868
ID AAV17868 standard; Protein; 183 AA.

AC AAV17868;

DT 20-AUG-1999 (first entry)

DE Streptococcus streptavidin.

KW Avidin; streptavidin; batroxobin; fibrinogen converting enzyme;
KM hybrid; fusion protein; sealant; surgery; reduce bleeding; fibrin.

OS Streptococcus sp.

PN WO929838-A1.

PD 17-JUN-1999.

PF 09-DEC-1998; 98WO-US26086.

PR 09-DEC-1997; 97US-0067978.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Cederholm-Williams SA;

XX WPI; 1999-385599/32.

DR N-PSDB; AAX80198.

XX A fibrinogen-converting enzyme fusion protein

PS Disclosure; Page 28; 35pp; English.

CC The present invention describes a fibrinogen-converting enzyme fusion
CC protein (FCE). The fusion protein is a multidomain protein comprising:
CC (a) a FCE; and (b) a first member of a binding pair that is linked to
CC the FCE chain; (i) directly by bonds utilizing the N-terminal amino
CC groups, the C-terminal carboxy groups or side-chain functionalities;
CC (ii) via a bifunctional linkage moiety linking the groups or
CC functionalities; or (iii) by the first member binding to the second
CC member of the binding pair, where the second member of the binding pair
CC is covalently attached to the first polypeptide chain. The FCE can be
CC used in a method for producing fibrin. Fibrin is useful as a sealant in
CC surgery to, e.g. reduce bleeding by sealing blood vessels, and tissues
CC that have been dissected either in surgery or through wounding. The
CC fusion protein allows for the removal of the fibrinogen converting
CC enzyme from the fibrin sealant preparation via the binding of
CC streptavidin to a biotin solid support. The present sequence represents
CC Streptococcus streptavidin as given in the present invention.

CC Sequence 183 AA;

Query Match 100.0%; Score 936; DB 20; Length 183;
Best Local Similarity 100.0%; Pred. No. 9.5e-76;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRIIVAAIAVSLTIVSITASASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
DB 1 MKRIIVAAIAVSLTIVSITASASADPSKDSKAQVSAEAGITGTWYNOLGSTFIYTAGAD 60
QY 61 GALTGTYESAVGNASERIVLTGRYDSAPATDGSCTALGWTVMKNNYNRAHSATTWSGOY 120
DB 61 GALTGTYESAVGNASERIVLTGRYDSAPATDGSCTALGWTVMKNNYNRAHSATTWSGOY 120
QY 121 VGGAERINTOMLITSGTTEANAMKSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDA 180
DB 121 VGGAERINTOMLITSGTTEANAMKSTLVGHDTFTKVPASASIDAKKAGVNNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183

RESULT 7

AAV44701
ID AAV44701 standard; Protein; 183 AA.

AC AAV44701;

DT 25-APR-2000 (first entry)

DE Streptavidin protein for recombinant PART27 vector.

KW Potato proteinase inhibitor-II; PPI-II; streptavidin; worm;

KM insect; plant-noxious protein; pest resistance; moth; insect; weevil;

KW grub; beetle; fly; thrip; locust; cricket; borer; mite; looper;

KW insecticidal.

OS Unidentified

PN WO200004049-A1.

PD 27-JAN-2000.

XX

PF 15-JUL-1999; 99MO-N200110.
 XX
 PR 15-JUL-1998; 98NZ-0331062.
 XX
 PA (HORT-) HORTICULTURE & FOOD RES INST NEW ZEALAND.
 XX
 PI Christeller JT, Sutherland PM, Murray C, Markwick NP, Philip BA;
 PI Malone LA, Burgess EPJ;
 XX
 DR WPI; 2000-171244/15.
 DR N-PSDB; AAC249867.
 XX
 PT New chimeric polypeptide and composition comprising the polypeptide
 PT useful for conferring pest resistance on plants -
 XX
 PS Disclosure; Fig 12; 11pp; English.
 XX
 CC The present sequence is streptavidin, a plant-noxious protein.
 CC Recombinant vector, PART27 expressing a chimeric polypeptide comprising
 CC streptavidin mature peptide fused to the potato proteinase inhibitor-II
 CC (PPI-II) signal peptide is targeted to the vacuole.
 CC Transformation of plant genome with the vector can produce pest
 CC resistance in plants, plant derived products and stored harvest
 CC material. Pests that can be controlled include, cotton bollworm,
 CC tropical army-worm, European corn-borer or red mite, tobacco horn worm,
 CC loopers, rice stem borer, porina, cutworms, diamondback moth, potato
 CC tuber moth, codling moth, indian meal moth, gypsy moth, argentine stem
 CC weevil, clover root weevil, grass-grubs, corn rootworm, rice and wheat
 CC weevils, mealworms, flour beetles, black field cricket, locusts,
 CC sawflies, western flower thrips, Hessian flies or two-spotted mite.
 XX
 SQ Sequence 183 AA;
 Query Match 100.0%; Score 936; DB 21; Length 183;
 Best Local Similarity 100.0%; Pred. No. 9.5e-76;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRKIIVAAIVSLTTSITASASADPSKDSKAOVSAEAGITGTWYNOLGSTFIYTAGAD 60
 DB 1 MRKIIVAAIVSLTTSITASASADPSKDSKAOVSAEAGITGTWYNOLGSTFIYTAGAD 60
 QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSGLTGVAMKNNYRNAHSATTWSGOY 120
 DB 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSGLTGVAMKNNYRNAHSATTWSGOY 120
 QY 121 VGGAEARINTOMLTSGTTEANAMKSTLVGHDTFTVKYKPSAASIDAAKAGVNNGNPLDA 180
 DB 121 VGGAEARINTOMLTSGTTEANAMKSTLVGHDTFTVKYKPSAASIDAAKAGVNNGNPLDA 180
 QY 181 VQQ 183
 DB 181 VQQ 183
 RESULT 8
 AAB30692 standard; Protein; 183 AA.
 ID AAB30692;
 AC AAB30692;
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Amino acid sequence of a streptavidin polypeptide.
 XX
 KM Streptavidin; tumour cell; cancer; adenocarcinoma;
 KM hematological malignancy.
 XX
 OS Streptomyces avidinii.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..24
 FT /note= "signal peptide"
 XX

PN WO200075333-A1.
 XX
 PD 14-DEC-2000.
 XX
 PF 05-JUN-2000; 2000MO-US15595.
 XX
 PR 07-JUN-1999; 99US-0137900.
 PR 03-DEC-1999; 99US-0168976.
 XX
 PA (NEOR-) NEORX CORP.
 XX
 XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;
 PI WPI; 2001-091213/10.
 DR N-PSDB; AAC66561.
 XX
 PT New vector constructs for expressing genomic streptavidin fusion
 PT proteins which are useful for targeting tumour cells associated with
 XX cancer, e.g. adenocarcinomas -
 XX
 PS Claim 18; Fig 4; 100pp; English.
 XX
 CC The present sequence represents a streptavidin polypeptide. The sequence
 CC is used to construct vectors of the invention. The specification
 CC describes vector constructs for expressing streptavidin fusion proteins.
 CC The vector comprises a first nucleic acid encoding genomic streptavidin
 CC or its functional variant operatively linked to a promoter, and a
 CC cloning site for insertion of a second nucleic acid sequence encoding a
 CC polypeptide to be fused with streptavidin. Inteposed between the
 CC promoter and the first nucleic acid sequence. Alternatively, the vector
 CC construct comprises a first nucleic acid, operatively linked to a
 CC promoter, encoding a polypeptide to be fused with streptavidin, and a
 CC cloning site for insertion of a second nucleic acid encoding at least
 CC 129 amino acids of streptavidin or its functional variant. The fusion
 CC proteins are useful for targeting tumour cells, particularly tumour cells
 CC associated with cancer, e.g. adenocarcinomas or hematological
 CC malignancies. The vector construct is useful for expressing of
 CC streptavidin fusion proteins. In particular, these are useful as tools
 CC for medical diagnostics and therapeutic purposes, e.g. for detecting the
 CC presence or absence of, or treating, a target site within a mammalian
 CC host.
 XX
 SQ Sequence 183 AA;
 Query Match 100.0%; Score 936; DB 22; Length 183;
 Best Local Similarity 100.0%; Pred. No. 9.5e-76;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRKIIVAAIVSLTTSITASASADPSKDSKAOVSAEAGITGTWYNOLGSTFIYTAGAD 60
 DB 1 MRKIIVAAIVSLTTSITASASADPSKDSKAOVSAEAGITGTWYNOLGSTFIYTAGAD 60
 QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSGLTGVAMKNNYRNAHSATTWSGOY 120
 DB 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGSGLTGVAMKNNYRNAHSATTWSGOY 120
 QY 121 VGGAEARINTOMLTSGTTEANAMKSTLVGHDTFTVKYKPSAASIDAAKAGVNNGNPLDA 180
 DB 121 VGGAEARINTOMLTSGTTEANAMKSTLVGHDTFTVKYKPSAASIDAAKAGVNNGNPLDA 180
 QY 181 VQQ 183
 DB 181 VQQ 183
 RESULT 9
 AAY80512 standard; Protein; 186 AA.
 ID AAY80512;
 AC AAY80512;
 XX
 DT 06-JUN-2000 (first entry)
 XX

DE Streptomyces avidinii sps protein.
 XX
 KW plant somatic tissue degeneration; plant essential factor; depletion;
 KM viability; sps gene; plant development; plant morphology; flower;
 KM fruit plant.
 OS Streptomyces avidinii.
 XX
 PN WO200007427-A2.
 XX
 PD 17-FEB-2000.
 XX
 PF 30-JUL-1999; 99WO-IL00420.
 XX
 PR 03-AUG-1998; 98IL-0125632.
 XX
 PA (AGRI-) AGRIC RES ORG.
 XX
 PI Kapulnik Y, Ginzberg I;
 XX
 DR WPI; 2000-195402/17.
 DR N-PSDB; AA291073.
 XX
 PT Degeneration of somatic plant tissue by expression of a heterologous
 PT protein, useful for controlling plant development and morphology, such
 PT as decreasing the number of flowers present to increase the number of
 PT fruit -
 XX
 PS Examples; Page 84; 91pp; English.
 XX
 CC The invention relates to a method of effecting degeneration of a somatic
 CC plant tissue by expressing a heterologous protein capable of binding a
 CC plant essential factor (PEF), in somatic plant tissue cells, where
 CC heterologous protein expression causes depletion of the PEF so the plant
 CC viability is maintained, while simultaneous degeneration of the somatic
 CC plant tissue is effected. This sequence represents the Streptomyces
 CC avidinii sps protein as an example of a heterologous protein introduced
 CC into the plants. The methods can provide for the selective and optionally
 CC reversible cell degeneration in somatic plant tissue. They can be used
 CC for artificially controlling plant development and morphology. They can
 CC be used e.g. to decrease the number of flowers in fruit producing plants
 CC so as to increase the number of fruits which reach maturity.
 CC
 XX
 SQ Sequence 186 AA;
 Query Match 100.0%; Score 936; DB 21; Length 186;
 Best Local Similarity 100.0%; Pred. No. 9.7e-76;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKRIVVAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNQIGSTFIYTAGAD 60
 Db 4 MKRIVVAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNQIGSTFIYTAGAD 63
 QY 61 GALTGTYSAVGNASSRYLTLGRYDSAPATDSSGALGWTAVWKNNYRAHSAATTWSSGY 120
 Db 64 GALTGTYSAVGNASSRYLTLGRYDSAPATDSSGALGWTAVWKNNYRAHSAATTWSSGY 123
 QY 121 VGGAARINTOWLTSGTTEANAMKSTLVGHDTFTKVPKSAASIDAACKAGVNNGNPLDA 180
 Db 124 VGGAARINTOWLTSGTTEANAMKSTLVGHDTFTKVPKSAASIDAACKAGVNNGNPLDA 183
 QY 181 VQO 183
 Db 184 VQO 186

XX
 DE Streptomyces avidinii mst protein.
 XX
 KW plant somatic tissue degeneration; plant essential factor; depletion;
 KM viability; mst gene; plant development; plant morphology; flower;
 KM fruit plant.
 OS Streptomyces avidinii.
 XX
 PN WO200007427-A2.
 XX
 PD 17-FEB-2000.
 XX
 PF 30-JUL-1999; 99WO-IL00420.
 XX
 PR 03-AUG-1998; 98IL-0125632.
 XX
 PA (AGRI-) AGRIC RES ORG.
 XX
 PI Kapulnik Y, Ginzberg I;
 XX
 DR WPI; 2000-195402/17.
 DR N-PSDB; AA291074.
 XX
 PT Degeneration of somatic plant tissue by expression of a heterologous
 PT protein, useful for controlling plant development and morphology, such
 PT as decreasing the number of flowers present to increase the number of
 PT fruit -
 XX
 PS Examples; Page 85; 91pp; English.
 XX
 CC The invention relates to a method of effecting degeneration of a somatic
 CC plant tissue by expressing a heterologous protein capable of binding a
 CC plant essential factor (PEF), in somatic plant tissue cells, where
 CC heterologous protein expression causes depletion of the PEF so the plant
 CC viability is maintained, while simultaneous degeneration of the somatic
 CC plant tissue is effected. This sequence represents the Streptomyces
 CC avidinii mst protein as an example of a heterologous protein introduced
 CC into the plants. The methods can provide for the selective and optionally
 CC reversible cell degeneration in somatic plant tissue. They can be used
 CC for artificially controlling plant development and morphology. They can
 CC be used e.g. to decrease the number of flowers in fruit producing plants
 CC so as to increase the number of fruits which reach maturity.
 CC
 XX
 SQ Sequence 184 AA;
 Query Match 99.5%; Score 911; DB 21; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.7e-75;
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RKIIVAAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNQIGSTFIYTAGAD 61
 Db 3 RKIIVAAALAVSLTIVSITASADPSKDSKAQVSAEAGITGTWYNQIGSTFIYTAGAD 62
 QY 62 ALTGTYSAVGNASSRYLTLGRYDSAPATDSSGALGWTAVWKNNYRAHSAATTWSSGY 121
 Db 63 ALTGTYSAVGNASSRYLTLGRYDSAPATDSSGALGWTAVWKNNYRAHSAATTWSSGY 122
 QY 122 GGAERINTOWLTSGTTEANAMKSTLVGHDTFTKVPKSAASIDAACKAGVNNGNPLDAV 181
 Db 123 GGAERINTOWLTSGTTEANAMKSTLVGHDTFTKVPKSAASIDAACKAGVNNGNPLDAV 182
 QY 182 QO 183
 Db 183 QO 184

RESULT 10
 ID AAY80513
 AC AAY80513 standard; Protein; 184 AA.
 XX
 DT 06-JUN-2000 (first entry)

RESULT 11
 ID AAP70492
 AC AAP70492 standard; protein; 182 AA.
 XX
 DT AAP70492;

DT 25-MAR-2003 (updated)
 DT 06-MAR-1991 (first entry)
 XX
 XX Streptavidin sequence.
 DE
 XX
 KM Streptavidin: N-terminal; fusion gene; fusion protein;
 XX
 OS Streptomyces avidinii.
 XX
 PN MO8705026-A.
 XX
 PD 27-AUG-1987.
 XX
 PF 24-FEB-1987; 87MO-US00397.
 XX
 PR 24-FEB-1986; 86US-0833324.
 XX
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 PA (UNY) UNIV NEW YORK STATE.
 XX
 PI Cantor CR, Axel R, Garana C;
 XX
 DR WPI; 1987-250198/35.
 DR N-PSDB; AAN70810.
 XX
 PT DNA encoding streptavidin - obtd. by restriction endo-nuclease
 PT digestion of chromosomal DNA of Streptomyces avidinii
 XX
 PS disclosure; Figure 3; 54pp; English.
 XX
 CC The sequence is that of streptavidin from S.avidinii. Streptavidin
 CC may be expressed from a fusion gene comprising its coding gene and DNA
 CC encoding a target protein of interest, where the streptavidin has
 CC binding sites for biotin or deriv. The streptavidin has 4 binding
 CC sites free for biotin, and is produced free of biotin
 CC contamination. Improved streptavidins may also be produced by
 CC site-directed mutagenesis. The fused gene may be used to produce
 CC labelled, chemically-modified proteins in vivo, and to isolate
 CC proteins when only the sequence of the gene is known.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 182 AA;
 Query Match 99.1%; Score 928; DB 8; Length 182;
 Best Local Similarity 99.5%; Pred. No. 4.9e-75;
 Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 RKIVAAIAVSLTTVSITASADPSKSKAQSAAEAGITGTWYNQLGSTFIVTAGADG 61
 DB 1 RKIVAAIAVSLTTVSITASADPSKSKRAQVSAEAGITGTWYNQLGSTFIVTAGADG 60
 QY 62 ALTGYESAVGNASERVLTGRYDSAPATDGSCTALGWTVAAMKNNRNNAHSATTWSGQYV 121
 DB 61 ALTGYESAVGNASERVLTGRYDSAPATDGSCTALGWTVAAMKNNRNNAHSATTWSGQYV 120
 QY 122 GGAERINTOWLTSCTTEANAMKSTLVGHDTFTKVPKSAASIDAAGVNNGNPLDAV 181
 DB 121 GGAERINTOWLTSCTTEANAMKSTLVGHDTFTKVPKSAASIDAAGVNNGNPLDAV 180
 QY 182 QQ 183
 DB 181 QQ 182

RESULT 12

AAV84020
 ID AAV84020 standard; protein; 183 AA.
 XX
 AC AAV84020;
 XX
 DT 03-JUL-2000 (first entry)
 XX
 DE Amino acid sequence of the wildtype streptavidin monomer.

XX
 KM Streptavidin; monomer; biotin binding domain; functional domain;
 KM biotin; cell adhesion peptide; adaptor; streptavidin/biotin interaction;
 KM substrate; vascular device; prosthesis.
 XX
 OS Streptomyces avidinii.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /note= "propeptide"
 XX
 PN WO200011152-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 25-AUG-1999; 99WO-US19481.
 XX
 PR 25-AUG-1998; 98US-0097816.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 PI Staceyton PS, McDevitt TC, Nelson KE;
 XX
 DR WPI; 2000-224689/19.
 XX
 PT Mutant streptavidin molecule, useful as an adaptor and a coating for
 PT substrates such as vascular devices or prostheses, comprises a biotin
 PT binding domain and a secondary functional domain -
 XX
 PS disclosure; Page 35-36; 43pp; English.
 XX
 CC The present sequence represents a wildtype streptavidin monomer. The
 CC specification describes streptavidin molecules comprising a biotin
 CC binding domain and a secondary functional domain. The molecules are
 CC adaptors with inherent effector function. They can therefore, bind to
 CC biotin and also have another function, for e.g. binding to a cell
 CC through the secondary functional domain comprising a cell adhesion
 CC peptide. The streptavidin molecules are useful as adaptors to bring,
 CC via a streptavidin/biotin interaction, the secondary functional domain
 CC into proximity with a cell or molecule to be affected and as a coating
 CC for substrates such as vascular devices or prostheses. Therefore, any
 CC compound of interest, such as a nucleic acid, protein, peptide, organic
 CC compound, inorganic compound, polysaccharide or a combination, can be
 CC targeted, delivered or immobilized using them. The molecules, when
 CC comprising an antibody for its secondary functional domain, are also
 CC useful in diagnostic applications for detecting analytes.
 XX
 SQ Sequence 183 AA;
 Query Match 99.1%; Score 928; DB 21; Length 183;
 Best Local Similarity 99.5%; Pred. No. 4.9e-75;
 Matches 182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MKRIVAAIAVSLTTVSITASADPSKSKAQSAAEAGITGTWYNQLGSTFIVTAGAD 60
 DB 1 MKRIVAAIAVSLTTVSITASADPSKSKAQSAAEAGITGTWYNQLGSTFIVTAGAD 60
 QY 61 GALTGTYESAVGNASERVLTGRYDSAPATDGSCTALGWTVAAMKNNRNNAHSATTWSGQY 120
 DB 61 GALTGTYESAVGNASERVLTGRYDSAPATDGSCTALGWTVAAMKNNRNNAHSATTWSGQY 120
 QY 121 VGGAERINTOWLTSCTTEANAMKSTLVGHDTFTKVPKSAASIDAAGVNNGNPLDA 180
 DB 121 VGGAERINTOWLTSCTTEANAMKSTLVGHDTFTKVPKSAASIDAAGVNNGNPLDA 180
 QY 181 VQQ 183
 DB 181 VQQ 183

RESULT 13

AAM59217
 ID AAM59217 standard; protein; 183 AA.

```

XX AC AAMS9217;
XX XX 27-AUG-1998 (first entry)
XX DT S. avidinii streptavidin mutant protein #1.
XX DE Streptavidin; ligand; binding affinity; mutant; isolation;
XX KW purification; recover; immobilise.
XX OS Streptomyces avidinii.
XX OS Synthetic.
XX OS
XX FT Key Location/Qualifiers
XX FT Misc-difference 68..71
XX FT /label= VTAR68ESAV
XX FT /note= "Wild type E8AV is replaced by VTAR. Numbering
XX FT is from the start of the mature protein"
XX FT
XX FT EP835934-AZ.
XX PN 15-APR-1998.
XX PD 09-OCT-1997; 97EP-0117504.
XX PF 10-OCT-1996; 96DE-1041876.
XX PR (BIOA-) INST BIOANALYTIK GMBH GOETTINGEN.
XX PA
XX PI Skerra A, Voss S;
XX PI
XX DR MPI; 1998-218868/20.
XX DR N-PSDB; AAV34715.
XX DR
XX PT Streptavidin mutants with higher binding affinity for peptide
XX PT ligands - have mutation in amino acid region 44-53, used to isolate,
XX PT purify or determine fusion proteins including these ligands
XX PT
XX PS Disclosure; Page -: 21pp; German.
XX PS
XX CC This sequence represents a mutant streptavidin protein isolated from
XX CC Streptomyces avidinii where the residues E8AV at position 44-47 of
XX CC the mature wild type sequence are replaced by VTAR. This sequence is
XX CC used to produce mutants which are used in a method to assay the binding
XX CC affinity of streptavidin mutants. These mutants have a mutation within
XX CC the amino acid (aa) region 44-53 of the wild-type protein show a higher
XX CC binding affinity than the wild-type for peptide ligands that include
XX CC the sequence of formula Trp-X-His-Pro-Gln-Phe-Y-Z where X = any aa; Y
XX CC and Z are both Gly, or Y = Glu and Z = Arg or Lys. Recombinant
XX CC streptavidin mutants can be used to isolate, purify and determine
XX CC proteins or to determine/recover substances that contain
XX CC streptavidin-binding groups. Such compounds may also be used to
XX CC immobilise fusions on microtitre plates, microbeads or sensor chips.
XX CC NOTE: This sequence does not appear in the specification but has
XX CC been constructed from the wild-type streptavidin sequence represented
XX CC in AAV34714.
XX CC
XX SQ Sequence 183 AA;
XX
XX Query Match 98.2%; Score 919; DB 19; Length 183;
XX Best Local Similarity 98.4%; Pred. No. 3,1e-74;
XX Matches 180; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 MKRIVVAALAVSLITTSITASASADPSKDSKAQVSAEAGITGTWNQSGFTIYAGAD 60
XX 1 MKRIVVAALAVSLITTSITASASADPSKDSKAQVSAEAGITGTWNQSGFTIYAGAD 60
XX DB 1 MKRIVVAALAVSLITTSITASASADPSKDSKAQVSAEAGITGTWNQSGFTIYAGAD 60
XX 61 GALTGTYESAVGNAESRYYVLTRGYDSAPALDGSGLALGWTAMKNNYRNAHSATTWSGQY 120
XX 61 GALTGTYYTARAGNAESRYYVLTRGYDSAPALDGSGLALGWTAMKNNYRNAHSATTWSGQY 120
XX QY 121 VGGAEARINTQWLLTSGTTEANAKWSTLVGHDTFTVKYKPSAASIDAKKAGVNNGNPLDA 180
XX 121 VGGAEARINTQWLLTSGTTEANAKWSTLVGHDTFTVKYKPSAASIDAKKAGVNNGNPLDA 180

```

```

DB 121 VGGAEARINTQWLLTSGTTEANAKWSTLVGHDTFTVKYKPSAASIDAKKAGVNNGNPLDA 180
QY 181 VQO 183
DB 181 VQO 183
XX
XX RESULT 14
XX AAMS9218
XX ID AAMS9218 standard; Protein, 183 AA.
XX AC AAMS9218;
XX XX 27-AUG-1998 (first entry)
XX DT S. avidinii streptavidin mutant protein #2.
XX DE Streptavidin; ligand; binding affinity; mutant; isolation;
XX KW purification; recover; immobilise.
XX OS Streptomyces avidinii.
XX OS Synthetic.
XX OS
XX FT Key Location/Qualifiers
XX FT Misc-difference 68..71
XX FT /label= IGAR68ESAV
XX FT /note= "Wild type E8AV is replaced by IGAR. Numbering
XX FT is from the start of the mature protein"
XX FT
XX FT EP835934-AZ.
XX PN 15-APR-1998.
XX PD 09-OCT-1997; 97EP-0117504.
XX PF 10-OCT-1996; 96DE-1041876.
XX PR (BIOA-) INST BIOANALYTIK GMBH GOETTINGEN.
XX PA
XX PI Skerra A, Voss S;
XX PI
XX DR MPI; 1998-218868/20.
XX DR N-PSDB; AAV34716.
XX DR
XX PT Streptavidin mutants with higher binding affinity for peptide
XX PT ligands - have mutation in amino acid region 44-53, used to isolate,
XX PT purify or determine fusion proteins including these ligands
XX PT
XX PS Disclosure; Page -: 21pp; German.
XX PS
XX CC This sequence represents a mutant streptavidin protein isolated from
XX CC Streptomyces avidinii where the residues E8AV at position 44-47 of
XX CC the mature wild type sequence are replaced by IGAR. This sequence is
XX CC used to produce mutants which are used in a method to assay the binding
XX CC affinity of streptavidin mutants. These mutants have a mutation within
XX CC the amino acid (aa) region 44-53 of the wild-type protein show a higher
XX CC binding affinity than the wild-type for peptide ligands that include
XX CC the sequence of formula Trp-X-His-Pro-Gln-Phe-Y-Z where X = any aa; Y
XX CC and Z are both Gly, or Y = Glu and Z = Arg or Lys. Recombinant
XX CC streptavidin mutants can be used to isolate, purify and determine
XX CC proteins or to determine/recover substances that contain
XX CC streptavidin-binding groups. Such compounds may also be used to
XX CC immobilise fusions on microtitre plates, microbeads or sensor chips.
XX CC NOTE: This sequence does not appear in the specification but has
XX CC been constructed from the wild-type streptavidin sequence represented
XX CC in AAV34714.
XX CC
XX SQ Sequence 183 AA;
XX
XX Query Match 98.0%; Score 917; DB 19; Length 183;
XX Best Local Similarity 98.4%; Pred. No. 4,7e-74;
XX Matches 180; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MRKIYVAALAVSLTIVSITASADSPSKQVSAEAGITGTWYNOLGSTFIYTAGAD 60
 DB 1 MRKIYVAALAVSLTIVSITASADSPSKQVSAEAGITGTWYNOLGSTFIYTAGAD 60
 QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGGTALGWTVMKNNYRNAHSATTWSGOY 120
 DB 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGGTALGWTVMKNNYRNAHSATTWSGOY 120
 QY 121 VGGAEARINTQWLITSGTTEANAMKSTLVGHDTFTVKKPSAASIDAKKAGVNNGNP 180
 DB 121 VGGAEARINTQWLITSGTTEANAMKSTLVGHDTFTVKKPSAASIDAKKAGVNNGNP 180
 QY 181 VQOQ 183
 DB 181 VQOQ 183
 RESULT 15
 ID AAY84021 standard; protein; 186 AA.
 XX AAY84021;
 AC AAY84021;
 XX
 DT 03-JUL-2000 (first entry)
 XX
 DE Amino acid sequence of a fibrinonectin/streptavidin fusion.
 XX
 KM Streptavidin; monomer; biotin binding domain; functional domain;
 KM biotin; cell adhesion peptide; adaptor; streptavidin/biotin interaction;
 KM substrate; vascular device; prosthesis; fibrinonectin.
 XX
 OS Synthetic.
 OS Streptomyces avidinii.
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /note= "propeptide"
 XX
 PN MO20001152-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 25-AUG-1999; 99MO-US19481.
 XX
 PR 25-AUG-1998; 98US-0097816.
 XX
 PA (UNIM) UNIV WASHINGTON.
 XX
 PI Stayton PS, McDevitt TC, Nelson KE;
 XX
 DR MPI; 2000-224689/19.
 XX
 PT Mutant streptavidin molecule, useful as an adaptor and a coating for
 PT substrates such as vascular devices or prostheses, comprises a biotin
 PT binding domain and a secondary functional domain -
 XX
 PS Example 4; Page 36-37; 43pp; English.
 CC The present sequence represents a fusion of fibrinonectin and
 CC streptavidin. It is a streptavidin molecule of the invention. The
 CC specification describes streptavidin molecules comprising a biotin
 CC binding domain and a secondary functional domain. The molecules are
 CC adaptors with inherent effector function. They can therefore, bind to
 CC biotin and also have another function, for e.g. binding to a cell
 CC through the secondary functional domain comprising a cell adhesion
 CC peptide. The streptavidin molecules are useful as adaptors to bring,
 CC via a streptavidin/biotin interaction, the secondary functional domain
 CC into proximity with a cell or molecule to be affected and as a coating
 CC for substrates such as vascular devices or prostheses. Therefore, any
 CC compound of interest, such as a nucleic acid, protein, peptide, organic
 CC compound, inorganic compound, polysaccharide or a combination, can be
 CC targeted, delivered or immobilized using them. The molecules, when
 CC comprising an antibody for its secondary functional domain, are also

CC useful in diagnostic applications for detecting analytes.
 XX

SO Sequence 186 AA.

Query Match 96.3%; Score 901.5; DB 21; Length 186;
 Best Local Similarity 96.2%; Pred. No. 1.2e-72;

Matches 179; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MRKIYVAALAVSLTIVSITASADSPSKQVSAEAGITGTWYNOLGSTFIYTAGAD 60
 DB 1 MRKIYVAALAVSLTIVSITASADSPSKQVSAEAGITGTWYNOLGSTFIYTAGAD 60
 QY 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGGTALGWTVMKNNYRNAHSATTWS 117
 DB 61 GALTGTYESAVGNAESRYVLTGRYDSAPATDGGTALGWTVMKNNYRNAHSATTWS 120
 QY 118 GQYVGGAEARINTQWLITSGTTEANAMKSTLVGHDTFTVKKPSAASIDAKKAGVNNGNP 177
 DB 121 GQYVGGAEARINTQWLITSGTTEANAMKSTLVGHDTFTVKKPSAASIDAKKAGVNNGNP 180
 QY 178 LDAVQO 183
 DB 181 LDAVQO 186

Search completed: October 27, 2003, 10:44:00
 Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 27, 2003, 22:14:12 ; Search time 1931 Seconds
(without alignments)
2303.324 Million cell updates/sec

Title: US-09-589-870B-2
Perfect score: 936
Sequence: 1 MKRIYVAIAVSLTVSIFA.....IDAKKAGVNGNPLDAVQO 183

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09589870/runac_27102003_104446_16625/app_query.fasta_1.327
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORX=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09589870.0CCN_1_1_2810@runac_27102003_104446_16625 -NCPU=6 -ICPU=3
-NO_MAP -LANG=QUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gse_hum.*
18: em_gse_inv.*
19: em_gse_pin.*
20: em_gse_vrt.*
21: em_gse_fun.*
22: em_gse_mam.*
23: em_gse_mus.*
24: em_gse_pro.*
25: em_gse_rtd.*
26: em_gse_png.*
27: em_gse_vr1.*
28: gb_gse1.*

29: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	165.5	17.7	435	12	BM080994
2	165.5	17.7	534	12	BM080994
3	162.5	17.4	598	12	BM183382
4	150.5	16.1	537	9	AL919330
5	150.5	16.0	406	12	BM531019
6	148	15.8	570	13	BU409120
7	148	15.8	578	13	BU426481
8	148	15.8	612	13	BU422332
9	147	15.7	538	12	BM489973
10	146.5	15.7	542	12	BI979837
11	146.5	15.7	608	13	EX263336
12	146	15.6	576	13	BU447569
13	145.5	15.5	622	13	BU293374
14	145	15.5	575	13	BU447766
15	145	15.5	582	13	BU408647
16	145	15.5	582	13	EX263637
17	145	15.5	577	13	BU385250
18	143.5	15.3	495	12	BM529407
19	143	15.3	601	14	CS508889
20	143	15.3	654	14	CS507800
21	141	15.1	568	14	CS507500
22	139.5	14.9	473	12	BU038085
23	132.5	14.2	588	13	EX081587
24	129.5	13.8	430	12	BI708243
25	129	13.6	548	14	CS507054
26	127.5	13.6	592	12	BI672746
27	126	13.5	593	13	BU418594
28	124.5	13.3	809	13	BU386325
29	124.5	13.3	901	13	BU382470
30	124	13.2	771	13	BU383814
31	122.5	13.1	495	9	AM632946
32	122.5	13.1	505	13	EX264326
33	120.5	12.9	379	10	BG414084
34	115	12.3	614	13	BU410176
35	113.5	12.1	470	12	BI709652
36	112	12.0	427	13	EX265638
37	110.5	11.8	342	14	CS490982
38	110.5	11.8	401	14	CS488823
39	109	11.6	453	13	EX278781
40	108.5	11.6	346	14	CS486748
41	108.5	11.6	352	14	CS487730
42	108.5	11.6	367	14	CS491403
43	108.5	11.6	376	14	CS486669
44	108.5	11.6	381	14	CS490090
45	108.5	11.6	384	14	CS487378

ALIGNMENTS

RESULT 1
LOCUS BM080994
DEFINITION BM080994 435 bp mRNA linear EST 14-NOV-2001
IMAGE:5159603.5' similar to SW:AVID_CHICK P02701 AVIDIN PRECURSOR.
[1] ; mRNA sequence.
ACCESSION BM080994
VERSION BM080994.1 GI:16927924
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes

REFERENCE
AUTHORS

1 Cyprinidae; Danio.
 (bases 1 to 435)
 Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE
JOURNAL
COMMENT

Washu Zebrafish EST Project 1998
 Unpublished
 Other_Estis: ft78g06.x1
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@wustl.edu
 The library was constructed by Dr. Z. Gong. DNA Sequencing by:
 Washington University Genome Sequencing Center St. Louis. Please
 contact Zhiyuan Gong for further information on this library
 (National University of Singapore, Department of Biological
 Sciences, Lower Kent Ridge Road, Singapore 119260).
 High quality sequence stop: 393.
 Location/Qualifiers

FEATURES
source

1..435
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:5159603"
 /sex="female"
 /dev_stage="4-5 month"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Gong zebrafish ovary"
 /note="Organ: ovary (pooled); Vector: pBluescript SK-;
 Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolated from
 the ovaries of 2 female adult zebrafish (4-5 month old).
 cDNAs were made using oligo-dT primers and inserted into
 lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
 mass-excised to pBluescript SK- following the Washington
 University protocol
 (http://genome.wustl.edu/est/lambda_protocol.shtml).
 Please contact Zhiyuan Gong for further information on
 this library (National University of Singapore,
 Department of Biological Sciences, Lower Kent Ridge Road,
 Singapore 119260)."
 BASE COUNT 99 a 100 c 131 g 105 t
 ORIGIN

Alignment Scores:

Pred. No.: 7.8e-08 Length: 435
 Score: 165.50 Matches: 54
 Percent Similarity: 48.80% Conservative: 27
 Best Local Similarity: 32.53% Mismatches: 50
 Query Match: 17.68% Indels: 35
 DB: 12 Gaps: 10

US-09-589-870b-2 (1-183) x BM080994 (1-435)

QY 18 IIEThrAlaSerAlaSerAlaSProSerLySaSPSerLyAlaGlnValSerAla 37
 Db 3 ACTCTAAAGACGACGCGCCGCAACGAGCTCTCAACGCCCAATGCAAGGTGAGCTCC-- 59
 QY 38 GUAAlaGlyTlLeThrGlyThrTPrTyAsGlnLeuGlySerThrPheIleValThrala 57
 Db 60 ---TGTAAATGTCACCGGCTTTGGCCGCAATGAGCTTGCTTACGCTGCGGCTGAAGGCG 116
 QY 58 ---GlyAlaAspGlyAlaLeuThrGlyThrTyGluSerAlaValGlyAsnAlaGluSer 76
 Db 117 GAAGGCTCGAG-----CTCAGAGGTGTTTACGACGACGCGGCTT-----GAGAGT 161
 QY 77 ArgTyValLeuThrGlyArgTyArgSerAla-----ProIaThrAspGly 92
 Db 162 ACGCAGGAGCGCGGATCATCATCGCTCCGCCGCGCATCATCGGATGTGTAGTACGGG 221

QY 93 SerGlyThrAlaLeuGlyTPrThrValAlaTPrLySaSPSerLyArgAsnAlaHisSer 112
 Db 222 ACTCAGCCTACCGCTCTCTTCCGTGCTATGGAGAAAGT-----TCG 266
 QY 113 AlaThrThrTPrSerGlyGlyTVal-----GlyGlyAlaGlnAlaArgIleAsn 129
 Db 267 TGCCTGCTGCTGGGTGGTCAGTGCCTTATTATTAGATGACGGGACACAGTCTGAAA 323
 QY 130 ThrGlnTPrLeuLeuThrSer---GlyThrThrGlnAlaAsnAlaTPrLySerThrLeu 148
 Db 324 ACCCTTGAGATGTTGCCGAGCTTGACACAACTTTGGCAATGCTCGGGAAGACACCGA 383
 QY 149 ValGlyHisAspTrpPheThrLyValLySProSerAlaAlaSerIleAspAlaAlaLys 168
 Db 384 ATGGAGAAAGATATTTTTTC----- 404
 QY 169 LysAlaGlyValaAsn 174
 Db 405 AAGACTGAGCTTCAAAT 422

RESULT 2
LOCUS
DEFINITION

B1980006/c 534 bp mRNA linear EST 24-OCT-2001
 ft78g06.x1 Gong zebrafish ovary Danio rerio cDNA clone
 IMAGE:5159603 3' similar to SW:AVID_CHICK P02701 AVIDIN PRECURSOR.
 [1] ;, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

B1980006
 B1980006.1 GI:16367292
 EST.
 Danio rerio (zebrafish)
 Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Osteichthys; Cypriniformes
 ; Cyprinidae; Danio.
 1 (bases 1 to 534)

REFERENCE
AUTHORS

Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 Washu Zebrafish EST Project 1998
 Unpublished
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@wustl.edu
 The library was constructed by Dr. Z. Gong. DNA Sequencing by:
 Washington University Genome Sequencing Center St. Louis. Please
 contact Zhiyuan Gong for further information on this library
 (National University of Singapore, Department of Biological
 Sciences, Lower Kent Ridge Road, Singapore 119260).
 Seq primer: 17 from Gidco
 High quality sequence stop: 421.
 Location/Qualifiers

FEATURES
source

1..534
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:5159603"
 /sex="female"
 /dev_stage="4-5 month"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Gong zebrafish ovary"
 /note="Organ: ovary (pooled); Vector: pBluescript SK-;
 Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolated from
 the ovaries of 2 female adult zebrafish (4-5 month old).
 cDNAs were made using oligo-dT primers and inserted into
 lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
 mass-excised to pBluescript SK- following the Washington


```

Oy 108 ArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyVal-----GlyGlyAla 124
Db 313 -----TCGAGCTCAGAGTGGCTGGTGGTCAAGTCTTATTATTAGATACAGCGGGA 360
Oy 125 GluAlaArgIleAsnThrGlnTrpLeuLeuThrSer---GlyThrThrGluAlaAsnAla 143
Db 361 CAAGTG---CTGAAACCTTCTGATGTTGCCAGAGTTTGCAGACAATTGGCAAGTCC 417
Oy 144 TrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSerAlaIleSer 163
Db 418 TGGGAGAGACCAAGATGGAGAGATATTATTTTTC----- 453
Oy 164 IleAspAlaAlaLysLysAlaGlyValaAsn 174
Db 454 -----AAGACTGAGATTCAAT 471

RESULT 4
AL919330/c 537 bp mRNA linear EST 18-SEP-2002
LOCUS AL919330 PUR-Z1+22 Danio rerio cDNA clone 061-G04-2, mRNA sequence.
DEFINITION AL919330
ACCESSION AL919330
VERSION AL919330.1 GI:23184628
KEYWORDS EST
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 537)
Lee,S., Ruan,H., Ma,W., Xu,M., Lo,J., He,Y., Liu,F., Eun,A., Wen,Z.
and Peng,U.
15,000 unique zebrafish EST clusters from two cDNA libraries
Contact: Peng J
Unpublished
JOURNAL
COMMENT Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengj@imcb.a-star.edu.sg
Clone requests: pengj@imcb.a-star.edu.sg.
Location/Qualifiers
FEATURES
source
1..537
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="061-G04-2"
/tissue_type="whole embryo or fish"
/dev_stage="mixed stages"
/clone_1lb="PUR-Z1+22"
BASE COUNT 145 a 146 c 107 g 139 t
ORIGIN
US-09-589-870b-2 (1-183) x AL919330 (1-537)

Alignment Scores:
Pred. No.: 414e-06 Length: 537
Score: 150.50 Matches: 49
Percent Similarity: 48.68% Conservative: 25
Best Local Similarity: 32.24% Mismatches: 43
Query Match: 16.08% Indels: 35
DB: 9 Gaps: 10

US-09-589-870b-2 (1-183) x AL919330 (1-537)
Oy 32 AlaGlnValSerAlaGlnAlaGlyIleThrGlyThrTrpTyraSngInLeuGlySer 51
Db 527 GCAAGGAGAGCTCC-----TGTATGTCACCGGTGTTGGCACATGAGCTTGCTCT 474
Oy 52 ThrPheIleValThrAla---GlyAlaAspGlyAlaLeuThrGlyThrTyGluSerAla 70
Db 473 ACGCTGCGGCTGAAGCGGAGGCTCGAG-----GTCAGAGCTGTTTACCAGACCGCG 420
Oy 71 ValGlnAsnAlaGluSerArgTyValLeuThrGlyArgTyraSpsrAla----- 87

```

```

Db 419 GTT-----GAGAGTACCGACGAGCCGACAGTATACCCGCTCCGCCGATCATC 369
Oy 88 ---ProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThrValAlaTrpLysAsnAsn 106
Db 368 GGGATGTCAGTACGAGGAGCTACCGCTACCGCTCTTTCTTCGTCGTCATGAGGAGAAAGT 309
Oy 107 TyrArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyVal-----GlyGly 123
Db 308 -----TCGCTCTGCTGGTGTGTCAGTCTTATTATTAGATGACCGG 264
Oy 124 AlaGlnAlaArgIleAsnThrGlnTrpLeuLeuThrSer---GlyThrThrGluAlaAsn 142
Db 263 GCACAAAGTA---CTGAAACCTTCTGATGTTGCCAGAGTTTGCAGACAATTGGCAAGT 207
Oy 143 AlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSerAlaAla 162
Db 206 GCTTGGGAGACACACAGATGGAGAGATATTTTTC----- 168
Oy 163 SerIleAspAlaAlaLysLysAlaGlyValaAsn 174
Db 167 -----AAGACTGAGATTCAAT 150

RESULT 5
BM531019 406 bp mRNA linear EST 19-FEB-2002
LOCUS BM531019
DEFINITION fW9JF04.Y1 Gong zebrafish ovary Danio rerio cDNA clone
IMAGE:5618455.5, similar to SW:AVID_CHICK P02701 AVIDIN PRECURSOR.
[1] : mRNA sequence.
ACCESSION BM531019
VERSION BM531019.1 GI:18739320
KEYWORDS EST
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 406)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Maria,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Treising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck,R., Ritzer,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCam,R., Waterston,R.
and Wilson,R.
Washu Zebrafish EST Project 1998
Unpublished
JOURNAL
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zdrafish@wustl.edu
The library was constructed by Dr. Z. Gong. DNA sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
Seq primer: T3 ET from Amersham
High quality sequence stop: 395.
Location/Qualifiers
FEATURES
source
1..406
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5618455"
/sex="female"
/dev_stage="4-5 month"
/lab_host="DH108 (phage-resistant)"
/clone_1lb="Gong zebrafish ovary"
/note="Organ: ovary (pooled); Vector: pBluescript SK-;
Site 1: XhoI; Site 2: EcoRI; Poly A+ RNA was isolated from
the ovaries of 2 female adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo

```

mass-excised to pBluescript SK- following the Washington University protocol (http://genome.wustl.edu/esf/lambda_protocol.shtml). Please contact Zhiyuan Gong for further information on this library (National University of Singapore, Department of Biological Sciences, Lower Kent Ridge Road, Singapore 119260).

BASE COUNT 88 a 95 c 125 g 98 t

ALIGNMENT SCORES:

Pred. No.: 3,16e-06 Length: 406
Score: 150.00 Matches: 48
Percent Similarity: 50.68% Conservative: 26
Best Local Similarity: 32.88% Mismatches: 50
Query Match: 16.03% Indels: 22
DB: 12 Gaps: 9

US-09-589-870b-2 (1-183) x BMS31019 (1-406)

```

QY 18 ILEThAlaSerAlaSerAlaProSerLyAspSerLyAlaGlnValSerAlaAla 37
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 8 ATCTCAAAAGCAGCGTCCGCAAGAGTCTCAACCGCAATGCAAGTGAGCTCC-- 64

QY 38 GLUAlaGlyIleThrGlyThrTPYrAsnGlnLeuGlySerThrPheIleValThrAla 57
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 65 ---TGTAAATCTACCGGTGTTGGCCGAATGAGCTTGCTCTACGCTCGCGGAGGAGGCC 121

QY 58 ---GlyAlaAspGlyAlaLeuThrGlyThrTyGluSerAlaValGlyAsnAlaGluSer 76
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 122 GAAGGCTCGGAG-----GTCAAGAGGTGTACCAAAACCGC-----GGTGAAGAT 166

QY 77 ArgTyValLeuThrGlyArgTyArgPserAla-----ProAlaThrAspGly 92
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 167 ACTCAGGAGAGCCCGCATCATCACCGCTCCGCCGATCATCGGATGTCAGTGAACGGG 226

QY 93 SerGlyThrAlaLeuGlyThrThrValAlaIleThrLeuAsnThrTyArgAsnAlaHisSer 112
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 227 ACTCAGGCTACCGCTCTCTTCTCCGCTGCTAGTGGAGAAAGT-----TCG 271

QY 113 AlaThrThrTyPserGlyGlnTyVal-----GlyGlyAlaGluAlaArgIleAsn 129
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 272 TGTCTGCTGCTGGGTGGTGTGCTGCTTATTATTAGATGAGCGGCACAAAGT---CTGAA 328

QY 130 ThrGlnIleuLeuLeuThrSer---GlyThrThrGluAlaAsnAlaIleuLeu 148
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 329 ACCTTCTGATGTGTGGCGACGTGTTCATACATTTGGCAAGTCTGGGGAGACACACAGA 388

QY 149 ValGlyHisAspThrPhe 154
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 389 ATGGAGAACATATTTT 406

RESULT 6
BU409120 570 bp mRNA linear EST 29-NOV-2002
LOCUS 603158328F1 CSFORBL03 Gallus gallus cDNA clone CHEST1767 5', mRNA
DEFINITION
sequence.
ACCESSION BU409120 GI:25901791
VERSION BU409120.1
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 570)
REFERENCE
AUTHORS Boardman,P.B., Sanz-Eguero,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard

```

Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST
)
PO Box 88, Manchester, M60 10D, UK
Tel: 01612008930
Fax: 01612160409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

```

source
1..570
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="CHEST1767"
/sex="Male and female"
/tissue_type="Chondrocytes isolated from growth plate cartilage"
/dev_stage="adult"
/lab_host="PH10B"
/clone_lib="CSFORBL03"
/notes="Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-timed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BspI and BamHI sites (5'ggccggcgtgcagccgagccgcgaaaaag) (5'aattcttttcgggagtcgcggcgagccgcgcagc)"
BASE COUNT 146 a 169 c 143 g 112 t
ORIGIN
Alignment Scores:
Pred. No.: 8.32e-06 Length: 570
Score: 148.00 Matches: 48
Percent Similarity: 46.26% Conservative: 20
Best Local Similarity: 32.65% Mismatches: 59
Query Match: 15.81% Indels: 20
DB: 13 Gaps: 8

US-09-589-870b-2 (1-183) x BU409120 (1-570)

QY 34 ValSerAlaAlaGluAlaGlyIleThrGlyThrTPYrAsnGlnLeuGlySerThrPhe 53
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 100 CTCCTGCGCAAGAAAGTCTCGCTGACGTGGAAATGACCAAGATCTGGCTTCAACATG 159

QY 54 IleValIleThrAla---GlyAlaAspGlyAlaLeuThrGlyThrTyGluSerAlaValGly 72
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 160 ACCATCGGGGCTGTGAACAGCAAGAGTGAATTCACAGGACCTTACACCAACCGGTA--- 216

QY 73 AsnAlaGluSerArgTyValLeuThrGlyArgTyArgPserAlaProAlaThrAspGly 92
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 217 ACAGCCACATCAATGATGATCAAGAGTCAACACCATGATGGAGCAACAAACCATCAAC 276

QY 93 SerGlyThr-----AlaLeuGlyThrThrValAlaIleuLeuAsnThrTyArgAsnAla 110
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 277 AAGAGACCCACCCACCTTGGCTTCACTGTCATATGGAG-----TTTTC 324

QY 111 HisSerAlaThrThrTyPserGlyGlnTyVal-----GlyGlyAlaGluAlaArg 127
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 325 GAGTCACACACTGTCTTCAAGGCGGCAAGCTTCATAGACAGAAAGCGAAGAGCTC--- 381

QY 128 IleAsnThrGlnIleuLeuLeuThrSerGlyThrThrGlu---AlaAsnAlaIleuLeu 146
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 382 CTGAAGACCATGTGCTGCTGGCGTCAAGTGTAAATGACATGATGATGACTGAAGACT 441

QY 147 ThrIleValGlyHisAspThrPheThrTyValIleu-----Pro 159
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 442 ACCAGGCTGGCATCAACATCTTCACTGCTCGCCGACACAGAGAGTGAAGATGAGCCC 501

QY 160 SerAlaIleSerIleAspAla 166
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

Db 502 CGCAAGCCAGCAACATGCC 522

RESULT 7
BU426481 578 bp mRNA linear EST 29-NOV-2002
LOCUS 603958787f1 CSEQRBNO9 Gallus gallus cDNA clone CHEST930n22 5', mRNA
DEFINITION sequence.

ACCESSION BU426481
VERSION BU426481.1 GI:25919157
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 578)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.D.,
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source

1..578
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="CHEST930n22"
/sex="Male and female"
/tissue_type="Chondrocytes isolated from growth plate cartilage"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBNO9"
/note="Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., FNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 152 a 171 c 145 g 110 t

ORIGIN

Alignment Scores:
Pred. No.: 8.48e-06 Length: 578
Score: 148.00 Matches: 48
Percent Similarity: 46.26% Conservative: 20
Best Local Similarity: 32.65% Mismatches: 59
Query Match: 15.81% Indels: 20
DB: 13 Gaps: 8

US-09-589-870b-2 (1-183) x BU426481 (1-578)

QY 34 valserAlaIaagIaIaGlyIleThrGlyThrTyrTrpAsnGlnLeuGlySerThrPhe 53
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 97 CTCTTCGCAGAAAGTCTGCTGACTGAGAAATGACCAACATCTGGGCTCCAAACATG 156

QY 54 lIleValThrAla---GlyAlaAspGlyAlaLeuThrGlyThrTyrGlySerAlaValGly 72
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 157 ACCATCGGGGCTGTGAACCAAGGTGAATTCAACAGCAACCTTACACAGCGGTA--- 213
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

QY 73 AsnAlaGlySerArgTyrValIleuThrGlyArgTyrAspSerAlaProAlaThrAspGly 92
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 214 ACAGCCACATCAATGATGATCAAGTCAAGTCAACCTGATGATGAGACAAACACCATCAAC 273
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

QY 93 SerGlyThr-----AlaLeuGlyTyrThrValAlaThrTyrAsnAsnThrArgAsnAla 110
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 274 AAGAGACCCAGCCACCTTGGCTTCACTGCTCATTTGAAG-----TTTTC 321
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

QY 111 HisSerAlaThrThrTyrSerGlyGlnTyrVal-----GlyGlyAlaGlyAlaArg 127
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 322 GAGTCCACCATCTGCTTTCACGGCCAGGCTTTCATAGACAGAAACGGGAAGAGGTC-- 378
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

QY 128 lIeAsnThrGlnTyrPleuThrSerGlyThrThrGlu---AlaAsnAlaThrTyrSer 146
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 379 CTGAAGACCATGTGGCTGCTCGGCTCAAGTGTTAATGACATTTGGTGAAGTGAAGACT 438
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

QY 147 ThrIleuValGlyHisAspThrPheThrTyrValTyr-----Pro 159
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 439 ACCAGGCTGGCATCAACATCTTCACTGCTCGCCACACAGAAAGACTGATGGCC 498
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

QY 160 SerAlaAlaSerIleAspAla 166
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 499 CGCAAGCCAGCAACATGCC 519

RESULT 8
BU422332 612 bp mRNA linear EST 29-NOV-2002
LOCUS 603232324f1 CSEQRBNO9 Gallus gallus cDNA clone CHEST229p5 5', mRNA
DEFINITION sequence.

ACCESSION BU422332
VERSION BU422332.1 GI:25915008
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 612)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.D.,
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source

1..612
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone_lib="CSEQRBNO9"
/clone="CHEST229p5"
/sex="Male and female"
/tissue_type="Chondrocytes isolated from growth plate cartilage"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBNO9"
/note="Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated

using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldi et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

```
BASE COUNT      171 a      176 c      150 g      115 t
ORIGIN
Alignment Scores:
Pred. No.:      9,19e-06      Length:      612
Score:          148.00      Matches:      48
Percent Similarity: 46.26%      Conservative: 20
Best Local Similarity: 32.65%      Mismatches: 59
Query Match:      15.81%      Indels:      20
DB:              13      Gaps:      8

US-09-589-870B-2 (1-183) x BU422332 (1-612)
OY 34 ValSerAlaIaGluAlaGlyIleThrGlyThrTyrAsnGlnLeuGlySerThrPhe 53
    ::::::::::: ::::::::::: ||| :::::::::::
DB 99 CTCTCTGCCAGAAAGCTCGCTGACGTGGAAATGACCAACATCTGGGCTTCAACATG 158
    ::::::::::: ::::::::::: ||| :::::::::::
OY 54 IleValThrAla---GlyAlaAspGlyAlaLeuThrGlyThrTyrGlnSerAlaValGly 72
    ::::::::::: ::::::::::: ||| :::::::::::
DB 159 ACCATCGGGCGCTGTAACAGCAAGGTAATTCACAGGACCTACACAGCCGTA--- 215
    ::::::::::: ::::::::::: ||| :::::::::::
OY 73 AsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAspGly 92
    ::::::::::: ::::::::::: ||| :::::::::::
DB 216 ACAGCCACATCAATGATGATCAAGAGTCAACCTGATGAGACACAAACACCATCAAC 275
    ::::::::::: ::::::::::: ||| :::::::::::
OY 93 SerGlyThr-----AlaLeuGlyTyrThrValAlaIleThrLysAsnAsnTyrArgAsnAla 110
    ||| ::::::::::: ||| :::::::::::
DB 276 AAGAGACCCAGCCACCTTGGCTTCTCATCTGCAATTGGAG-----TTTTCa 323
    ||| ::::::::::: ||| :::::::::::
OY 111 HisSerAlaIleThrTyrSerGlyGlnTyrVal-----GlyGlyAlaGluAlaArg 127
    ::::::::::: ::::::::::: ||| :::::::::::
DB 324 GAGTCCACCACTGCTTCCACGGGCAAGTCTTCATACAGAAAGGAGAGAGTC--- 380
    ::::::::::: ::::::::::: ||| :::::::::::
OY 128 IleAsnThrGlnTyrLeuLeuThrSerGlyThrThrGlu---AlaAsnAlaTyrLysSer 146
    ::::::::::: ::::::::::: ||| :::::::::::
DB 381 CTGAAGCCATGCTGGCTGCTGCTGCAAGTGAATGATGATGCTGATGAGAAAGCT 440
    ::::::::::: ::::::::::: ||| :::::::::::
OY 147 ThrLeuValGlyHisAspThrPheThrLysValLys-----Pro 159
    ||| ||| ::::::::::: ||| :::::::::::
DB 441 ACCAGGCTCGGCATCAACATCTTCACTCGCTCGGCACACAGAAAGATGAGATGGCC 500
    ||| ||| ::::::::::: ||| :::::::::::
OY 160 SerAlaIleSerIleAspAla 166
    ||| ||| ::::::::::: ||| :::::::::::
DB 501 CGCAAGCCAGCAACATGCC 521
    ||| ||| ::::::::::: ||| :::::::::::

RESULT 9      538 bp      mRNA      linear      EST 07-FEB-2002
LOCUS      BM489973
DEFINITION      Bp2n.pK002.a9 Normalized Chicken Pituitary/Hypothalamus/Pineal
                Library (Bp2n) Gallus gallus cDNA clone Bp2n.pK002.a9 5' similar
                to emb|CAC34569.1 (A0111647) avadin [Gallus gallus], mRNA sequence.
ACCESSION      BM489973
VERSION      BM489973.1 GI:18610904
KEYWORDS      EST.
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE      Porter,T.E. and Cogburn,L.A.
AUTHORS      1 (bases 1 to 538)
TITLE      ESTs from Normalized Chicken Pituitary/Hypothalamus/Pineal cDNA
```

Library, USDA/IFAPS Animal Genome Project
Unpublished
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu www.chickest.udel.edu.
Location/Qualifiers
1. 538
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Commercial broiler chickens"
/db_xref="taxon:9031"
/clone="Bp2n.pK002.a9"
/sex="Male and Female"
/tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"
/dev_stage="Embryonic (d12,d14,d19) ; post-hatch (1,3,5,7,9 weeks)"
/lab_host="E. coli EMD10B"
/clone_lib="Normalized Chicken
Pituitary/Hypothalamus/Pineal Library (Bp2n)"
/note="Vector: PCWVSP06; Library made from equivalent
pools of total RNA isolated from each tissue at different
ages. Single pass sequencing from 5'-end"

```
BASE COUNT      140 a      157 c      137 g      104 t
ORIGIN
Alignment Scores:
Pred. No.:      9.8e-06      Length:      538
Score:          147.00      Matches:      48
Percent Similarity: 46.26%      Conservative: 20
Best Local Similarity: 32.65%      Mismatches: 59
Query Match:      15.71%      Indels:      20
DB:              12      Gaps:      8

US-09-589-870B-2 (1-183) x BM489973 (1-538)
OY 34 ValSerAlaIaGluAlaGlyIleThrGlyThrTyrAsnGlnLeuGlySerThrPhe 53
    ::::::::::: ::::::::::: ||| :::::::::::
DB 63 CTCTCTGCCAGAAAGCTCGCTGACGTGGAAATGAGACACGACCTCGGCTTCAACATG 122
    ::::::::::: ::::::::::: ||| :::::::::::
OY 54 IleValThrAla---GlyAlaAspGlyAlaLeuThrGlyThrTyrGlnSerAlaValGly 72
    ::::::::::: ::::::::::: ||| :::::::::::
DB 123 ACCATCGGGCGCTGTAACAGCAAGGTAATTCACAGGACCTACCAAGCCGTA--- 179
    ::::::::::: ::::::::::: ||| :::::::::::
OY 73 AsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAspGly 92
    ::::::::::: ::::::::::: ||| :::::::::::
DB 180 ACAGCCACATCAATGATGATCAAGAGTCAACGATGATGAGACCAAAACCATCAAC 239
    ::::::::::: ::::::::::: ||| :::::::::::
OY 93 SerGlyThr-----AlaLeuGlyTyrThrValAlaIleThrLysAsnAsnTyrArgAsnAla 110
    ||| ::::::::::: ||| :::::::::::
DB 240 AAGAGACCCAGCCACCTTGGCTTCTCATCTGCAATTGGAG-----TTTTCa 287
    ||| ::::::::::: ||| :::::::::::
OY 111 HisSerAlaIleThrTyrSerGlyGlnTyrVal-----GlyGlyAlaGluAlaArg 127
    ::::::::::: ::::::::::: ||| :::::::::::
DB 288 GAGTCCACCACTGCTTCCACGGGCAAGTCTTCATACAGAAAGGAGAGTC--- 344
    ::::::::::: ::::::::::: ||| :::::::::::
OY 128 IleAsnThrGlnTyrLeuLeuThrSerGlyThrThrGlu---AlaAsnAlaTyrLysSer 146
    ::::::::::: ::::::::::: ||| :::::::::::
DB 345 CTGAAGACCATGCTGCTGCTGCTCAAGTGAATGATGATGATGATGATGATGATGATGAT 404
    ::::::::::: ::::::::::: ||| :::::::::::
OY 147 ThrLeuValGlyHisAspThrPheThrLysValLys-----Pro 159
    ||| ||| ::::::::::: ||| :::::::::::
DB 405 ACCAGGCTCGGCATCAACATCTTCACTCGCTCGGCACACAGAAAGATGAGATGGCC 464
    ||| ||| ::::::::::: ||| :::::::::::
OY 160 SerAlaIleSerIleAspAla 166
    ||| ||| ::::::::::: ||| :::::::::::
DB 465 CCAGAAAGCCAGCAACATGCC 485
    ||| ||| ::::::::::: ||| :::::::::::

RESULT 10      542 bp      mRNA      linear      EST 24-OCT-2001
LOCUS      B1979837
```

DEFINITION ft76d11.y1 Gong zebrafish ovary Danio rerio cDNA clone
IMAGE:5159109.5, similar to SW:AVID_CHICK P02701 AVIDIN PRECURSOR.
[1] ; mRNA sequence.
ACCESSION BI979837.1 GI:16366965
VERSION BI979837
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 542)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
S., Hillier,L., Kucba,T., Martin,J., Beck,C., Wyllie,T., Underwood
,K., Septoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
TITLE Washu Zebrafish EST Project 1998
JOURNAL Unpublished
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
The library was constructed by Dr. Z. Gong. DNA sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
Seq primer: T3 ET from Amersham
High quality sequence stop: 395.
LOCATION/Qualifiers
FEATURES
source
1..542
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5159109"
/sex="female"
/dev_stage="4-5 month"
/lab_host="PH10B (phage-resistant)"
/clone_lib="Gong zebrafish ovary"
/note="Organ: ovary (pooled); Vector: pBluescript SK-;
Site 1: XhoI; Site 2: EcoRI; Poly A+ RNA was isolated from
the ovaries of 2 female adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
mass-excisid to pBluescript SK- following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260).
BASE COUNT 142 a 105 c 148 g 147 t
ORIGIN
Alignment Scores: 1.12e-05 Length: 542
Pred. No.: 146.50 Matches: 49
Score: 47.68% Conservative: 23
Percent Similarity: 32.45% Mismatches: 44
Best Local Similarity: 15.65% Indels: 35
Query Match: 12 Gaps: 10
DB: 12
US-09-589-870b-2 (1-183) x BI979837 (1-542)
QY 33 GlnValSerAlaIaGuaIaGlyIleThrGlyThrTPrTyAsnGlnIeuGlySerThr 52
|||||
DB 5 CAGGTAGCTCC-----TGTATGTACACCGGTTTGGCGCAATAGCTTGCTCTACG 58
QY 53 PheIleValThrAla---GlyAlaAspGlyAlaIeuThrGlyThrTyrGluSerAlaVal 71

DB 59 CTCGGCGTGAAGCGGAGAGCTCGAG-----CTCAGAGCTGTTTCCACAGCCGGTT 112
|||
QY 72 GlyAsnAlaGluSerArgTyrValIeuThrGlyArgTyrAspSerAla----- 87
|||||
DB 113 -----GAGAGTACGACGAGCGGACCGAGGTATCATCCGCTCGCCGATCATCGG 163
|||||
QY 88 ProAlaThrAspGlySerGlyThrAlaIeuGlyTPrThrValAlaTPrTyAsnAsnTyr 107
|||||
DB 164 ATGCTCAGTACGCGGACCTACCGCTCTCTTCCTCCGCTATGGAGAAAGT--- 220
|||||
QY 108 ArgAsnAlaHisSerAlaThrThrTyrSerGlyGlnTyrVal-----GlyGlyAla 124
|||||
DB 221 -----TCGTGCTCGCTTGCGTGGTGCAGTCTTATTTAGATGACGGGCA 268
|||||
QY 125 GluAlaArgIleAsnThrGlnTPrIleuThrSer---GlyThrThrGluAlaAsnAla 143
|||||
DB 269 CAGGTG---CTGAATCCTTCTGATGTTCGCGCAGTCTTGACAGCAATTTGGCAAGTCC 325
|||||
QY 144 TPrTySerThrIleuValGlyHisAspThrPhrThrIlyValIysProSerAlaIaSer 163
|||||
DB 326 TCGGAGACACACAGATGAGAGAGATTTTTC----- 361
|||||
QY 164 IleAspAlaAlaIlyValAlaGlyValAsnAsn 174
|||||
DB 362 -----AGACTGAGATTCAAT 379
|||||
RESULT 11
BX263336 608 bp mRNA linear EST 27-FEB-2003
BX263336
LOCUS BX263336 AGENAE Gallus gallus multi-tissues normalized library
DEFINITION (gcag) Gallus gallus cDNA clone gcag0016c.g.24 5prim, mRNA
sequence.
ACCESSION BX263336
VERSION BX263336.1 GI:28585934
KEYWORDS EST
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 608)
Herault,F., Le Menth-Meizinger,V., Desert,C., Retout,E., Piumi,F.,
Klopp,C. and Douaire,M.
TITLE Construction and primary characterization of chicken normalized
multi-tissue cDNA libraries
JOURNAL Unpublished
COMMENT Contact: Douaire M
INRA, UMR INRA-ENSAR Genetique Animale
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
Tel: +33 (0) 2.23.48.54.63
Fax: +33 (0) 2.23.48.54.70
Email: Madeleine.Douaire@roazon.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
FEATURES
source
1..608
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gcag0016c.g.24"
/tissue_type="multi-tissues"
/tissue="from embryos to adults"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Gallus gallus multi-tissues normalized
library (gcag)"
/note="Vector: pTTT3D-pac; tissues: brain, embryos, kidney
adipose tissue, granulosa, utero-vaginal gland, oviduct,
small follicle, ovary, hypothalamus, pituitary gland,

ileon, jejunum, caecum, duodenum, spleen, fabricius gland
bone marrow, thymus, hematopoietic progenitor cells.
Clone distribution : AGENAE Resource centre, Francois
PUMI, Francois Pumi, intra.fr, INRA, CEA Radiobiologie et
Etude du genome (LREB), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, FRANCE"

BASE COUNT 111 a 204 c 179 g 114 t

ALIGNMENT SCORES:

Pred. No.: 1,31e-05 Length: 608
Score: 146.50 Matches: 50
Percent Similarity: 42.46% Conservative: 26
Best Local Similarity: 27.93% Mismatches: 62
Query Match: 15.65% Indels: 41
DB: 13 Gaps: 9

US-09-589-870B-2 (1-183) x BX263336 (1-608)

QY 34 ValSerAlaAlaGluAlaGlyIleThrGlyThrTyrPheAenGlnLeuGlySerThrPhe 53
DB 80 GTGGCTCTCCGGAGAGTCGAGCTGCAAGGCTCTGTGAGAGACGAGCTCGGCTCAACATG 139
QY 54 IleValThrAla---GlyAlaSerGlyAlaLeuThrGlyThrTyrGluSerAlaValGly 72
DB 140 ACCATCTGACCTGATGCGGAGCTTCTCGGCTCTTACGACAGCGGTG--- 196
QY 73 AsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAla-----ProAla 89
DB 197 ACAGCCACCAACAGAGATCTGTGTCTACCCCTGAAGGGCCAGACCCCGCGC 256
QY 90 ThrAspGlySerGlyThrAlaLeuGlyThrPheValAlaTyrPheAenThrArgAan 109
DB 257 ACCAAGGGGAGAGAGCCACCTTGGCTTCTACTGTGAGTGG-----CAGTTT 304
QY 110 AlaHisSerAlaThrThrTyrSerGlyGln---TyrValGly-----GlyAlaGluAla 126
DB 305 GCAGACTCCACCACTGCTTGTGGAGACAGTCTGTGAGACCGCGTGGAGAGAG--- 361
QY 127 ArgIleAenThrGlnThrLeuLeuThrSerGlyThr---ThrGluAlaAenAlaTyrPhe 145
DB 362 ATCTGTGAGATGCGATGTGTGCTGCGGAGAGAGTTCCTCCGCAAGGACACCTGGAAG 421
QY 146 SerThrLeuValGlyHisAspThrPheThrValValys----- 158
DB 422 GCCACGAGGTGACCAATGCTTCAACCGGTGCAAGTATGGGAGACATGCTGCGC 481
QY 158 ----- 158
DB 482 CCTCGGATCTCTGTGCAAGCGCTGGCTCTGTCTGCTCTCACACCGTCTCTGCTGC 541
QY 159 ProSerAlaAlaSerIleAspAlaAlaLeuAlaGlyValAenAenGlyAenPro 177
DB 542 CCACGTGCTCATCCGAGTGTGCTGTGACCAAGAGACCAATAAACCCTCGACCC 598
RESULT 12
BU447569 576 bp mRNA linear EST 29-NOV-2002
LOCUS BU447569 603764104F1 CSEQRBN13 Gallus gallus CDNA clone Chest687g2 5', mRNA
DEFINITION
sequence.
ACCESSION BU447569
VERSION BU447569.1 GI:25936880
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
AUTHORS 1 (bases 1 to 576)
TITLE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken CDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE

2235534
PUBMED 12445392
CONTACT: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..576

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Layer"

/db_xref="taxon:9031"

/clone="Chest687g2"

/sex="Female"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="CSEQRBN13"

/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldi et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

BASE COUNT 150 a 169 c 145 g 112 t

ORIGIN

Alignment Scores:
Pred. No.: 1.38e-05 Length: 576
Score: 146.00 Matches: 48
Percent Similarity: 46.94% Conservative: 21
Best Local Similarity: 32.65% Mismatches: 58
Query Match: 15.60% Indels: 20
DB: 13 Gaps: 8

US-09-589-870B-2 (1-183) x BU447569 (1-576)

QY 34 ValSerAlaAlaGluAlaGlyIleThrGlyThrTyrPheAenGlnLeuGlySerThrPhe 53
DB 95 CTCTCTGCCAGAAAGTCTGCTGACGTGGAAATGAGACCAAGATCTGGGCTCAACATG 154
QY 54 IleValThrAla---GlyAlaSerGlyAlaLeuThrGlyThrTyrGluSerAlaValGly 72
DB 155 ACCATCGGGGCTGTGAACAGCAGAGTGAATTCACAGGACCTTACATCACACCGTA 211
QY 73 AsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAspGly 92
DB 212 ACAGCCACATCAAAAGATCAAGATCAACACCATGATGAGACCAAAACCACTCAAC 271
QY 93 SerGlyThr-----AlaLeuGlyThrPheThrValAlaTyrPheAenAenThrArgAan 110
DB 272 AATAGACCCACGCCACTTGTGGCTTCAACCGTCAATTGGAG-----TTTCA 319
QY 111 HisSerAlaThrThrTyrSerGlyGlnTyrVal-----GlyAlaGluAlaArg 127
DB 320 GAGTCCACACATGCTTCAACGCGGCGGCTTCAATAGACAGAAATGGAGAGGTG--- 376
QY 128 IleAenThrGlnThrLeuLeuThrSerGlyThrGlu---AlaAsnAlaTyrPheSer 146
DB 377 CTGAAGACCATGTGCTGTGGGTCAAGTGAATGACATTGATGATGATGATGATGAAAGCT 436
QY 147 ThrLeuValGlyHisAspThrPheThrValValys-----Pro 159


```

/1ab_host="DH10B"
/clone_1ib="CSEORBN3"
/Note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methyated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
BASE COUNT      148 a      170 c      146 g      111 t
ORIGIN

Alignment Scores:
Pred. No.:      1,76e-05      Length:      575
Score:          145.00      Matches:      48
Percent Similarity: 46.26%      Conservative: 20
Best Local Similarity: 32.65%      Mismatches: 59
Query Match:     15.49%      Indels:      20
DB:              13      Gaps:      8

US-09-589-870B-2 (1-183) x BU447766 (1-575)
QY      34 valserAlaAlaGluAlaGlyIleThrGlyThrTyraAsnGlnLeuGlySerThrphe 53
      100 cttctGCGAGAAAGGCTGCTGCTGAGCAATGAGCAACGATCTGGCTCAACAG 159
QY      54 llevalThrAla---GlyAlaAspGlyAlaLeuThrGlyThrTyrgLuserAlaValGly 72
      160 ACCATCGGGGCTGTGAACAGCAGGATGCAATTCACAGCGCCTACATCACAGCCGTA--- 216
DB      73 AenAlaGuserArGlyrValLeuThrGlyArGlyrAspSerAlaProAlaThrAspGly 92
      217 ACAGCCACATCAATAGATCAAGAGTCCACCTGCATGGACACACAAACACATCAAC 276
QY      93 SerGlyThr-----AlaLeuGlyTrpThrValAlaTrpLysAsnAsnTyraAsnAla 110
      277 AAGAGAGCCAGCCACCTTGGCTTCAACGCTCAATTGGAAG-----TTTCA 324
DB      111 HisSerAlaThrThrTrpSerGlyGlnTyraVal-----GlyGlyAlaGluAlaArg 127
      325 GAGTCCACCACTGCTTCAAGCGCCAGTCTTCATAGACAGAAATGGGAAGAGTGC--- 381
QY      128 lleAsnThrGlnTrpLeuLeuThrSerGlyThrThrglu---AlaAsnAlaTrpLysSer 146
      382 CTGAAGACCTGTGGCTGCTGCTGCTCAAGTGAATGACTGCTGATGACTGGAAGACT 441
DB      147 ThrLeuValGlyHisAspThrThrThrLysValLys-----Pro 159
      442 ACCAGAGTGGGATCAACATCTTCACTCGCTCGACACAGAAAGAGTGAAGATGCC 501
QY      160 SerAlaAlaSerIleAsnAla 166
      502 CGCAAGCCAGCAACATGCC 522
DB

RESULT 15
BU408647      576 bp      mRNA      linear      EST 29-NOV-2002
LOCUS      603158704F1 CSEORBL03 Gallus gallus cDNA clone Chest177g1 5', mRNA
DEFINITION
sequence.
ACCESSION      BU408647
VERSION      BU408647.1 GI:25901318
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Gallus gallus
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

```

```

Phasianidae; Gallus.
1 (bases 1 to 576)
AUTHORS
Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Boech, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
TITLE
A Comprehensive Collection of Chicken cDNAs
JOURNAL
Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE
22335534
PUBMED
12445392
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 0161208930
Fax: 01612260409
Email: Simon.Hubbard@umist.ac.uk.
FEATURES
source
1..576
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="Chest177g1"
/sex="Male and female"
/tissue_type="Chondrocytes isolated from growth plate cartilage"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="CSEORBL03"
/Note="Vector: pBluescript II KS(+); Site_1: EcoRI;
Site_2: NotI; Modification of pBluescript II KS(+)
(Stratagene) vector to accommodate cDNA produced with the
Trimmed protocol (construction of uni-directionally
cloned cDNA libraries from messenger RNA for improved 3'
end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387
,624). Cut pBluescript II KS(+) with NotI and EcoRI.
Ligate in double stranded adaptor containing BamHI and
BamHI sites [5'ggcgcggtgagccgcgagccggaataaag]
[5'aattcttttcgagctcgggctcgacgc]"
BASE COUNT      148 a      170 c      147 g      111 t
ORIGIN

Alignment Scores:
Pred. No.:      1,76e-05      Length:      576
Score:          145.00      Matches:      48
Percent Similarity: 46.26%      Conservative: 20
Best Local Similarity: 32.65%      Mismatches: 59
Query Match:     15.49%      Indels:      20
DB:              13      Gaps:      8

US-09-589-870B-2 (1-183) x BU408647 (1-576)
QY      34 valserAlaAlaGluAlaGlyIleThrGlyThrTyraAsnGlnLeuGlySerThrphe 53
      95 cttctGCGAGAAAGGCTGCTGCTGAGCAATGAGCAACGATCTGGCTCAACAG 154
DB      54 llevalThrAla---GlyAlaAspGlyAlaLeuThrGlyThrTyrgLuserAlaValGly 72
      155 ACCATCGGGGCTGTGAACAGCAGGATGCAATTCACAGCGCCTACATCACAGCCGTA--- 211
QY      73 AenAlaGuserArGlyrValLeuThrGlyArGlyrAspSerAlaProAlaThrAspGly 92
      212 ACAGCCACATCAATAGATCAAGAGTCCACCTGCATGGACACACAAACACATCAAC 271
DB      93 SerGlyThr-----AlaLeuGlyTrpThrValAlaTrpLysAsnAsnTyraAsnAla 110
      272 AAGAGAGCCAGCCACCTTGGCTTCAACGCTCAATTGGAAG-----TTTCA 319
QY      111 HisSerAlaThrThrTrpSerGlyGlnTyraVal-----GlyGlyAlaGluAlaArg 127
      320 GAGTCCACCACTGCTTCAAGCGCCAGTCTTCATAGACAGAAATGGGAAGAGTGC--- 376
DB      128 lleAsnThrGlnTrpLeuLeuThrSerGlyThrThrglu---AlaAsnAlaTrpLysSer 146

```

```

Db      377 CTGAGACCACTGTGCTGCTGCGCAAGTGTATGACATGTGTGATGACTGGAAAGCT 436
Oy      147 ThrlauValGlyHisAspThrPheThrLysValLys-----Pro 159
Db      437 ACCAGGTCGGCATCAACATCTTCACTGCGCACACAGAGAGAGTGAAGATGGCC 496
Oy      160 SerAlaIaSerIleAspAla 166
Db      497 CGCAAGCCGACACATGCC 517

```

Search completed: October 28, 2003, 02:05:33
 Job time : 1939 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 28, 2003, 00:19:43 ; Search time 253 Seconds
(without alignments)
1952.558 Million cell updates/sec

Title: US-09-589-870B-2

Perfect score: 936
Sequence: 1 MRKIVVAIAIVSLTIVSITA.....IDAAKAGVNGNPNLDVQO 183

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlh
-Q=/cg22_1/USPTO.epool/US05989870/runat_27102003_104445_16575/app.query.fasta_1.327
-DB=N_Geneseq_19Jun03 -QFMT=fastad -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US05989870.GCEN_1_1.312.0/runat_27102003_104445_16575 -NCRU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -OSBLOCK=100 -LONLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	936	100.0	552	20	AAK80198	Streptococcus stre
2	936	100.0	563	21	AAZ91073	Streptomyces avidi
3	936	100.0	638	18	AAT73193	Wild-type streptav
4	936	100.0	638	19	AAV34714	S. avidinii strept
5	936	100.0	638	21	AAZ49867	Streptavidin gene
6	936	100.0	638	22	AAC86561	DNA encoding a str
7	936	100.0	1131	7	AAK60626	Sequence of a port
8	931	99.5	564	21	AAZ91074	Streptomyces avidi
9	927	99.0	552	14	AAQ35412	Streptavidin gene.
10	919	98.2	638	19	AAV34715	S. avidinii strept
11	917	98.0	638	19	AAV34716	S. avidinii strept
12	855.5	91.4	1239	22	AAC86563	DNA encoding a fus
13	854.5	91.3	1612	22	AAC86562	DNA encoding a fus
14	847	90.5	1280	22	AAC86564	Anti-CD20 single c
15	846	90.4	1356	15	AAQ70659	Scfv PRAS108 and p
16	838	89.5	495	21	AAZ91076	Streptomyces avidi
17	835	89.2	492	21	AAZ91075	Streptomyces avidi
18	835	89.2	498	9	AAK80186	Encodes biosynthe
19	834	89.1	498	18	AAT75491	DNA for streptavid
20	834	89.1	507	10	AAK90755	Sequence of strept
21	834	89.1	1266	24	ABA92275	Streptavidin-carbo
22	834	89.1	7306	25	ABZ77370	Nucleotide sequenc
23	834	89.1	18155	25	ABZ77369	Nucleotide sequenc
24	772	82.5	1881	25	AAQ50166	Biotin-combining s
25	745	79.6	1296	15	AAQ70660	Scfv PRAS109 and p
26	687.5	73.5	1173	18	AAT86312	Single-chain anti-
27	666	73.3	864	20	AAK89302	S. avidinii protein
28	682	72.9	1175	18	AAT86311	Single chain anti-
29	679	72.5	384	24	ABL61154	S. avidinii strept
30	679	72.5	401	21	AAZ49861	Streptavidin cDNA f
31	679	72.5	2019	17	AAT33850	DNA encoding strep
32	675	72.1	384	18	AAT73195	Recombinant Core-s
33	675	72.1	387	14	AAQ40718	Core streptavidin.
34	673.5	72.0	1257	15	AAQ70661	Scfv PRAS10 and p
35	672	71.8	626	21	AAZ49866	Potato proteinase
36	660	70.5	2055	17	AAT33851	DNA encoding mufan
37	658.5	70.4	1247	24	ABK12869	Chimeric SA-Past. D
38	656	70.1	1266	24	ABK12871	Chimeric TGF-B-SA
39	656	70.1	1329	24	ABK12872	Chimeric IL-2-SA D
40	656	70.1	1395	24	ABK12873	Chimeric IL-10-SA
41	656	70.1	1542	24	ABK12870	Chimeric B7.1-SA D
42	656	70.1	1566	24	ABK12874	Chimeric SA-CD40L
43	650	69.4	369	21	AAZ91077	Streptomyces avidi
44	650	69.4	376	21	AAZ91078	Streptomyces avidi
45	644	69.8	466	25	ABZ70370	Streptavidin-based

ALIGNMENTS

RESULT 1

ID AAK80198 standard; cDNA; 552 BP.

AC AAX80198;

DT 20-AUG-1999 (first entry)

XX Streptococcus streptavidin encoding cDNA.

DE Streptococcus streptavidin encoding cDNA.

XX Avidin; streptavidin; biotrobin; fibrinogen converting enzyme;

KW hybrid; fusion protein; sealant; surgery; reduce bleeding; fibrin; ss.

XX Streptococcus sp.

OS

XX

XX

PN

XX WO9929838-A1.

US-09-589-870B-2 (1-183) x AA291073 (1-563)

QY 1 MetArgLysIleValValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
 DB 10 ATGGCAAGATGCTGCTGGACGATCGCGCTTCCCTGACCAAGGCTCGATTACGGCC 69
 QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGluAlaGly 40
 DB 70 AGCGCTTGGCAACACCTCCCAAGACTCGAAGGCCAGGTCTCGGCGCGCGAGCCGGC 129
 QY 41 IleThrGlyThrTyrPheGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
 DB 130 ATCAGCGGACCTGTGACACCAAGCTCGCTGACCTTATCGTACCGCGGCGCGCAG 189
 QY 61 GlyAlaLeuThrGlyThrTyrGluSerAlaValAlaGlyAsnAlaGluSerArgTyrValLeu 80
 DB 190 GGGGCGCTGACCGGAACCTACGAGTCGGCGCGGCAACCGCCGAGAGCCGCTACGCTCG 249
 QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
 DB 250 ACCGGTCGTTACACAGCGCCCGGACCGGACCGGACCGGCGCGCTCGGTTGGAGC 309
 QY 101 ValAlaTyrLysAsnAsnTyrArgAsnAlaIleSerAlaThrThrTyrPheSerGlyGlnTyr 120
 DB 310 GTGGCGTGGAGAATACTACCGCAACGCCACTCCGCGACGAGCGGAGCGCGCATAC 369
 QY 121 ValGlyGlyAlaGluAlaArgIleAsnThrGlnTyrLeuLeuThrSerGlyThrThrGlu 140
 DB 370 GTGGCGGCGCGCGAGGAGATCAACACCCAGTGTGCTGTGACTCCGCGCACCGAG 429
 QY 141 AlaAsnAlaTyrLysSerThrLeuValGlyIleAspThrPheThrLysValLysProSer 160
 DB 430 GCCAAGCGCTGGAAAGTCTCAGCTGCTGCGGACGACCACTTCCCAAGGTGAAGCCGTCC 489
 QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
 DB 490 GCCGCTTCATCGACGCGGCGGAAGAGCGCGGTCAACAAGCGCAACCGCTCGAGCGCC 549
 QY 181 ValGlnGln 183
 DB 550 GTTCAGCAG 558

RESULT 3
 AAT73193
 ID AAT73193 standard; DNA; 638 BP.
 AC AAT73193;
 XX
 DT 27-APR-1998 (first entry)
 DE Wild-type streptavidin DNA.
 XX
 KM Streptavidin; biotin; anti-interference reagent; detection; muten;
 KM avidin; non-specific binding; ss.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT CDS 50..601
 FT /*tag= a
 FT /product= streptavidin
 FT sig_peptide 50..121
 FT /*tag= b
 FT mat_peptide 122..598
 FT /*tag= c
 FT /product= streptavidin
 XX
 PN DE19637718-Al.
 XX
 PD 02-OCT-1997.
 XX
 PF 16-SEP-1996; 96DE-1037718.

XX
 PR 01-APR-1996; 96DE-1013053.
 XX
 PA (BOER) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Brandstetter H, Deger A, Engh R, Kopetzki E, Mueller R;
 PI Schmidt U;
 DR WPI: 1997-482043/45.
 DR P-PSDB; AAM29306.
 XX
 PT Streptavidin and avidin muteins with reduced binding affinity for
 PT biotin - useful for reducing interference from nonspecific binding
 PT in assays
 XX
 PS Disclosure; Page 16-17; 26pp; German.
 XX
 CC This DNA sequence encodes a streptavidin which is used in a novel method
 CC of reducing interference from non-specific binding in assays. Muteins
 CC constructed from a core streptavidin or avidin sequence are selected that
 CC differ from the native polypeptide by at least one amino acid and have a
 CC binding affinity for biotin of less than 1010 l/mole. The biotin-bindable
 CC polypeptide may be present as a polymeric conjugate, e.g. with another
 CC polypeptide or protein, especially bovine serum albumin. These muteins
 CC are used as anti-interference reagents for reducing and/or avoiding
 CC nonspecific interactions in a process for detecting an analyte. In
 CC particular, they are used in assays where the streptavidin/avidin-biotin
 CC specific binding pair is involved for qualitative and/or quantitative
 CC determination of an analyte in a test sample, e.g. a heterogeneous
 CC immunoassay or a hybridization assay. Despite having a lower binding
 CC affinity for biotin, the muteins have high immunological cross-reactivity
 CC with native streptavidin and avidin.
 XX
 SQ Sequence 638 BP; 115 A; 244 C; 193 G; 86 T; 0 other;

Alignment Scores:
 Pred. No.: 5,93e-74 Length: 638
 Score: 936.00 Matches: 183
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-589-870B-2 (1-183) x AAT73193 (1-638)

QY 1 MetArgLysIleValValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
 DB 50 ATGGCAAGATGCTGCTGGACGATCGCGCTTCCCTGACCAAGGCTCGATTACGGCC 109
 QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGluAlaGly 40
 DB 110 AGCGCTTGGCAACACCTCCCAAGACTCGAAGGCCAGGTCTCGGCGCGCGAGCCGGC 169
 QY 41 IleThrGlyThrTyrPheGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
 DB 170 ATCAGCGGACCTGTGACACCAAGCTCGCTGACCTTATCGTACCGCGGCGCGCAG 229
 QY 61 GlyAlaLeuThrGlyThrTyrGluSerAlaValAlaGlyAsnAlaGluSerArgTyrValLeu 80
 DB 230 GGGGCGCTGACCGGAACCTACGAGTCGGCGCGGCAACCGCGAGAGCGCTACGCTCG 289
 QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
 DB 290 ACCGGTCGTTACACAGCGCCCGGACCGGACCGGACCGGCGCTCGGTTGGAGC 349
 QY 101 ValAlaTyrLysAsnAsnTyrArgAsnAlaIleSerAlaThrThrTyrPheSerGlyGlnTyr 120
 DB 350 GTGGCGTGGAGAATACTACCGCAACGCCACTCCGCGACGAGCGGCGCGCATAC 409
 QY 121 ValGlyGlyAlaGluAlaArgIleAsnThrGlnTyrLeuLeuThrSerGlyThrThrGlu 140
 DB 410 GTGGCGGCGCGGAGGAGATCAACACCCAGTGTGCTGTGACTCCGCGCACCGAG 469

QY 141 AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSer 160
 DB 470 GCCAACCCCTGGAAGTCCAGCTGTGGCCACGACCTTCACCAAGGTGAAGCCGTCC 529
 QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
 DB 530 GCCGCTTCATCGACGGCGGGAAGAGCCGGCGTCAACACGCAACCCGCTCGACGCC 589
 QY 181 ValGlnGln 183
 DB 590 GTTCAGCAG 598
 RESULT 4
 AAV34714
 ID AAV34714 standard; DNA; 638 BP.
 XX
 AC AAV34714;
 XX
 DT 27-AUG-1998 (first entry)
 XX
 DE S. avidinii streptavidin cDNA.
 XX
 KM Streptavidin; ligand; binding affinity; mutant; isolation;
 XX purification; recover; immobilise; ss.
 OS Streptomyces avidinii.
 XX
 FH Key Location/Qualifiers
 FT CDS 50..601
 FT FT /*tag= a
 FT sig_peptide 50..121
 FT FT /*tag= b
 FT mat_peptide 122..598
 FT FT /*tag= c
 FT /product= Streptavidin
 XX
 PN EP835934-A2.
 PD 15-APR-1998.
 XX
 PF 09-OCT-1997; 97EP-0117504.
 XX
 PR 10-OCT-1996; 96DE-1041876.
 XX
 PA (BIOA-) INST BIOANALYTIK GMBH GOETTINGEN.
 XX
 PI Skerra A, Voss S;
 XX
 DR WPI: 1998-218668/20.
 DR P-PSDB; AAMS9216.
 XX
 PT Streptavidin mutants with higher binding affinity for peptide
 PT ligands - have mutation in amino acid region 44-53, used to isolate,
 PT purify or determine fusion proteins including these ligands
 XX
 PS Disclosure; Page -: 21pp; German.
 XX
 CC This sequence encodes a wild-type streptavidin protein isolated from
 CC Streptomyces avidinii. This sequence is used to produce mutants which
 CC are used in a method to assay the binding affinity of streptavidin
 CC mutants. These mutants have a mutation within the amino acid (aa) region
 CC 44-53 of the wild-type protein show a higher binding affinity than the
 CC wild-type for peptide ligands that include the sequence of formula
 CC Trp-X-His-Pro-Gln-Phe-Y-Z where X = any aa; Y and Z are both Gly,
 CC or Y = Glu and Z = Arg or Lys. Recombinant streptavidin mutants can
 CC be used to isolate, purify and determine proteins or to
 CC determine/recover substances that contain streptavidin-binding groups.
 CC Such compounds may also be used to immobilise fusions on microtitre
 CC plates, microbeads or sensor chips.
 CC NOTE: This sequence does not appear in the specification but is used to
 CC make the mutant streptavidin sequence represented in AAV34715 and
 CC AAV34716.

SQ Sequence 638 BP; 115 A; 244 C; 193 G; 86 T; 0 other;
 Alignment Scores:
 Pred. No.: 5 93e-74 Length: 638
 Score: 936.00 Matches: 183
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0
 US-09-589-870B-2 (1-183) x AAV34714 (1-638)
 QY 1 MetArgLysIleValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
 DB 50 ATCGCGAAGATCGTGTTCGACGACCATCGCCGCTTCCTCGACCAAGGTCCGATTACGGCC 109
 QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGlnAglly 40
 DB 110 AGCGCTTCGGGAGACCCCTCCCAAGGACTCGAAGGCCCAAGTCTCGCCGCGAGGCCGCGC 169
 QY 41 IleThrGlyThrTrpTrpAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
 DB 170 ATCACCAGGACCTGCTGACACCAAGCTCGGCTCGACCTTCATCTGACCGGCGCCGAC 229
 QY 61 GlyAlaLeuThrGlyThrTrpGluSerAlaValGlyAsnAlaGluSerArgTyrValLeu 80
 DB 230 GCGCGCTCGACCGGAACTCAGAGTCCGCGCCGCGCAAGCGGAGCGGCTAGCTCTG 289
 QY 81 ThrGlyValGlyTrpAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
 DB 290 ACCGGTGTTCGACGAGGCGCCCGCCGACCGACCGGACCGGACCGCCCTCGTTGAGC 349
 QY 101 ValAlaTrpLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGlyValTyr 120
 DB 350 GTGGCTCGAAGAAATCACTACCGACGCCCACTCCGGACCACTGAGAGCGGCCAGTAC 409
 QY 121 ValGlyGlyAlaGlyAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThrGlu 140
 DB 410 GTCCGCGCGCGCGGAGGATCAACACCCAGTGGTGTGCTGACCTCCGCGCACCCAG 469
 QY 141 AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSer 160
 DB 470 GCCAACCCCTGGAAGTCCAGCTGTGGCCACGACCTTCACCAAGGTGAAGCCGTCC 529
 QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
 DB 530 GCCGCTTCATCGACGGCGGGAAGAGCCGGCGTCAACACGCAACCCGCTCGACGCC 589
 QY 181 ValGlnGln 183
 DB 590 GTTCAGCAG 598
 RESULT 5
 AAZ49867
 ID AAZ49867 standard; cDNA; 638 BP.
 XX
 AC AAZ49867;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Streptavidin gene for recombinant pART7 vector.
 XX
 KM Potato proteinase inhibitor-II; PPI-II; streptavidin; worm;
 KM insect; plant-noxious protein; pest resistance; moth; insect; weevil;
 KM grub; beetle; fly; thrip; locust; cricket; borer; mite; looper;
 XX insecticidal; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 50..601
 FT FT /*tag= a
 FT /product= "Streptavidin protein"

FT sig_peptide 50..121
 FT /*tag= b
 FT mat_peptide 122..598
 FT /*tag= c
 PN WO200004049-A1.
 XX 27-JAN-2000.
 PD 15-JUL-1999; 99WO-NZ00110.
 XX 15-JUL-1999; 98NZ-0331002.
 PR 15-JUL-1998; 98NZ-0331002.
 XX (HORT-) HORTICULTURE & FOOD RES INST NEW ZEALAND.
 PA Christaller JT, Sutherland PW, Murray C, Markwick NP, Phillip BA;
 PI Malone LA, Burgess EPJ;
 DR MPI; 2000-171244/15.
 DR P-PSDB; AAY44701.
 XX
 PT New chimeric polypeptide and composition comprising the polypeptide
 PT useful for conferring pest resistance on plants
 XX
 PS Disclosure; Fig 12; 11pp; English.
 CC The present sequence encodes streptavidin, a plant-noxious protein. This
 CC is used in the preparation of a recombinant vector, pART27 designed to
 CC express a chimeric polypeptide comprising streptavidin mature peptide
 CC fused to the potato proteinase inhibitor-II (PPI-II) signal peptide. The
 CC binary vector is targeted to the vacuole by PPI-II signal sequence.
 CC Transformation of plant genome with the vector can produce pest
 CC resistance in plants, plant derived products and stored harvest
 CC material. Pests that can be controlled include, cotton bollworm,
 CC tropical army-worm, European corn-borer or red mite, tobacco horn worm,
 CC loopers, rice stem borer, porina, cutworms, diamondback moth, potato
 CC tuber moth, codling moth, indian meal moth, gypsey moth, argentine stem
 CC weevil, clover root weevil, grass-grubs, corn rootworm, rice and wheat
 CC weevils, mealworms, flour beetles, black field cricket, locusts,
 CC sawflies, western flower thrips, Hessian flies or two-spotted mite.
 CC
 SQ Sequence 638 BP; 115 A; 244 C; 193 G; 86 T; 0 other;
 Alignment Scores:
 Pred. No.: 5.93e-74 Length: 638
 Score: 936.00 Matches: 183
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-589-870B-2 (1-183) x AA249867 (1-638)
 QY 1 MetArgLysIleValAlaAlaAlaIleAlaValSerLeuThrThyValSerIleThrAla 20
 DB 50 ATGCGCAAGATCGTCTGTCAGCCATCGCCGTTCCCTGACCGGTCTCGATTAGAGCC 109
 QY 21 SerAlaSerAlaAspProSerIleValAspSerIleValGlnValSerAlaIleAlaGlu 40
 DB 110 AGGCGTTCGGCAGACCCCTCCAGACTCGAAGGCCAGGTCTCGGCCCGCGAGCCGCG 169
 QY 41 ILeThpGlyThrThpIlyrAngInLeuGlySerThpHeIleValThraIagIyAlaAsp 60
 DB 170 ATTCACGGGACCGGTGACCAACGCTCGCTCATCTTATCTGACCGCGGCGCGAC 229
 QY 61 GAlaAlaLeuThrcIyThrTyGluSerAlaValAlaGlyAsnAlaGluSerArgTyValIleu 80
 DB 230 GGGGCGCTACCGGAACCTACGAGTGGCGCTGGCAACCGCCAGAGCCGCTACGTCCTG 289
 QY 81 ThnGlyArgTyAspSerAlaProAlaThraAspGlySerGlyThraIaleuGlyTyPThr 100
 DB 290 ACCGCTGTTACGACAGCCGCCCGGACGAGGAGCGAGCGAGCCGCCCTCGGTTGAGCG 349

QY 101 ValAlaTrpLysAsnAsnTyraArgAlaAlaHISerAlaLeuThrThrTropSerGlyGlnTy 120
 DB 350 GTGGCTTGAAATTAATTAACCGCAACGCCCACTCGGACCACTGAGCGGCCATGAC 409
 QY 121 ValGlyGlyAlaGlyAlaArgIleAsnThrcIyPleuLeuThrSerGlyThrcIy 140
 DB 410 GTGGGGGGGGCGGAGCGAGATCAACACCGAGTGGCTGCTGACCTCGGCGACCAACGAG 469
 QY 141 AlaAsnAlaTrpLysSerThrcIyValGlyHisAspThrPheThrcIyValLysProSer 160
 DB 470 GCCAAGCGCTGGAATCCACGCTGTCGGCCACGACACCTTACCAAGGTGAAGCGTTC 529
 QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAlaAsnGlyAsnProLeuAspAla 180
 DB 530 GCCGCTTCATGACGCGGCGGAGAAAGCGCGCTGACACAGGCAACCGCTCGAGCGC 589
 QY 181 ValGlnGln 183
 DB 590 GTTCAGCAG 598
 RESULT 6
 AAC86561
 ID AAC86561 standard; DNA; 638 BP.
 XX
 AC AAC86561;
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE DNA encoding a streptavidin polypeptide.
 XX
 KW Streptavidin; tumour cell; cancer; adenocarcinoma;
 XX hematological malignancy; ss.
 OS Streptomyces avidinii.
 XX
 FH Key Location/Qualifiers
 FT CDS 50..601
 FT /*tag= a
 FT /product= "streptavidin"
 FT sig_peptide 50..121
 FT /*tag= b
 PN WO200075333-A1.
 PD 14-DEC-2000.
 PF 05-JUN-2000; 2000WO-US15595.
 XX
 PR 07-JUN-1999; 99US-0137900.
 PR 03-DEC-1999; 99US-0168976.
 XX
 PA (NEOR-) NEORX CORP.
 PI
 XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JW;
 DR MPI; 2001-091213/10.
 DR P-PSDB; AAB30692.
 PT New vector constructs for expressing genomic streptavidin fusion
 PT proteins which are useful for targeting tumour cells associated with
 CC cancer, e.g. adenocarcinomas
 XX
 XX Claim 14; Fig 4; 100pp; English.
 CC The present sequence encodes a streptavidin polypeptide. The sequence
 CC is used to construct vectors of the invention. The specification
 CC describes vector constructs for expressing streptavidin fusion proteins.
 CC The vector comprises a first nucleic acid encoding genomic streptavidin
 CC or its functional variant operatively linked to a promoter, and a
 CC cloning site for insertion of a second nucleic acid sequence encoding a
 CC polypeptide to be fused with streptavidin, interspersed between the
 CC promoter and the first nucleic acid sequence. Alternatively, the vector
 CC construct comprises a first nucleic acid, operatively linked to a

CC promote encoding a polypeptide to be fused with streptavidin, and a
CC cloning site for insertion of a second nucleic acid encoding at least
CC 129 amino acids of streptavidin or its functional variant. The fusion
CC proteins are useful for targeting tumour cells, particularly tumour cells
CC associated with cancer, e.g. adenocarcinomas or hematological
CC malignancies. The vector construct is useful for expressing of
CC streptavidin fusion proteins. In particular, these are useful as tools
CC for medical diagnostics and therapeutic purposes, e.g. for detecting the
CC presence or absence of, or treating, a target site within a mammalian
CC host.

SQ Sequence 638 BP; 115 A; 244 C; 193 G; 86 T; 0 other;

Alignment Scores:

Pred. No.:	5.93e-74	Length:	638
Score:	936.00	Matches:	138
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-589-870B-2 (1-183) x AAC86561 (1-638)

OY	1	MetArgLysIleValIleAlaIleAlaIleAlaValSerLeuThrThrValSerIleThrAla	20
Db	50	ATGGCAAGATCGTCGTTGCAAGCCATCGCGTTTCCCTTACCAACGAGTCTCGATTACGGCC	105
OY	21	SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGlnAlaGly	40
Db	110	AGCCCTTCGGAGACCCCTCCAAAGACTCGAAGGCCCAAGGTCTCGGCGCCGAGCCGGC	165
OY	41	IleThrGlyThrTrpLysAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp	60
Db	170	ATCACCGGACCTGSTRCAACACAGCTGGCTGACTGACCTTATGTATGACCGCGGCGCCAC	225
OY	61	GlyAlaLeuThrGlyThrTyrgInSerAlaValGlyAsnAlaGlnSerArgTyValLeu	80
Db	230	GGCCGCTGACCGGAACCTACGAGTCGGCGCTGGCAACGCCGAGAGGCCCTACGTCCTG	285
OY	81	ThrGlyArgTyrrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrrThr	100
Db	290	ACGGGTGTTACACACAGCGCCCGGCGCACCGAGCGGACCGGACCCCTGTTTGACG	345
OY	101	ValAlaTrpLysAsnAsnTyrrArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyrr	120
Db	350	GTGGCTTGGAAGATACACTACCGCAACGCCCACTCGGACCAACGTGAGAGCGGCACTAC	405
OY	121	ValGlyGlyAlaGlnAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThrGln	140
Db	410	GTCCGCGCGCCCGAGGCGAGATCAACACCGAGTGGCTGACCTCGGACCAACCGAG	465
OY	141	AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSer	160
Db	470	GCCACGCGCTGGAAAGTCCACGCTGGTGGCCACGACACCTTCACCAAGTGAAGCCGCTC	525
OY	161	AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAlaAsnGlyAsnProLeuAspAla	180
Db	530	GCGCGCTTCATCGACCGCGGGAAGAGCGCGCTGCAACAGCGCAACCGCTCGACGCC	585
OY	181	ValGlnGln 183	
Db	590	GTTCAGCAG 598	
RESULT 7			
AAN60626			
ID AAN60626 standard; DNM; 1131 BP.			
XX AAN60626;			
XX			
DT 25-MAR-2003 (updated)			
DT 13-AUG-1991 (first entry)			
XX			
DE Sequence of a portion of SA307 which encodes a streptavidin-like			

DE polypeptide.
VY

KW Antibiotic; biotin binding affinity; fusion protein; ds.
 YY

OS Streptomyces
yy

FH	Key	Location/Qualifiers
ET	CDS	481 1000

/*tag= a
/not a=

CDS	482..1030
ET	/+33-
ET	p

```
ET /note= "AAP60624"
CDS 480 1031
```

$$/ * \text{Ea} = C$$

XX
F.T
XX
/note= "AAP60625"

PN
XX
XX
W08602077-A.

FD 10-APR-1986.
XX

01-001-1985; 85W0-0001901.

02-001-1984; 8405-06568/3.

PA (BIOJ) BIOGEN NV.
EA (MEAD/) MEADE H M.
EA (BIOJ) BIOGEN NV.

Meade HM, Garwin JL, Biogen NV:

WPI: 1986-106643/16.

L EODD; HAF000Z03; HAF00

XX
XX

streptavidin-like polypeptide, at

[illegible]

XX

steptavidin-like polypeptide (see

CC and a second sequence coding for

CC (Updated on 25-MAR-2003 to correct

SQ Sequence 1131 BP; 199 A; 409 C; 31

Alignment Scores:

Score: 936.00

Best Local Similarity: 100.00%

DB: 7

US-09-589-870B-2 (1-183) X AAN60626 (1-

Qy 1 MetArglySILEValValAlaAlaAla

Db 480 ATGCCGAGATCGTCGTTGCAGCC

QY 22 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGlnAlaGly 40
Db 540 ACCGCTTGGCGAGACCCCTCCAGAGACTCGAAGGCCAGGCTTCGCGCGCGAGGCCGCG 59
QY 41 IleThrGlyThnTriPtyrAsnGlnGluGlySerThrPheIleValThrAlaGlyAlaAsp 60
Db 600 ATCACC GGCACTGTGACACCAAGCTTCGGCTCGACCTTCATGCTGACCGCGGCGCCAC 65
QY 61 GlyAlaLeuThrGlyThnTriGluSerAlaValAlaGlyAsnAlaGluSerArgTyrrValLeu 80

Db 660 GGGCCCTGACGGAACTTACGATCGCGCCGCGCAACCGGAGAGCCGCTACGTCTTG 719
QY 81 ThrGlyArgThrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTTPThr 100
Db 720 ACCGGTGTGTTACAGACGGCCCGGCGGACGCGGAGCCGCCCTCGGTGGAGC 779
QY 101 ValAlaTrpIleAsnAntyTrpAsnAlaHisSerAlaThrThrTrpSerGlyIntYr 120
Db 780 GTGGCTGGAAAGATAACTACCCGCAACGCCCATCTCCGCGACAGTGGAGCGCCAGTAC 839
QY 121 ValGlyValAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrGlu 140
Db 840 GTGGCGGCGCCGAGGAGGATCAACACCCAGTGGCTGTGACTCCGCGACACCGAG 899
QY 141 AlaAsnAlaTrpIleSerThrLeuValGlyHisAspThrThrIleValIleProSer 160
Db 900 GCCAAGCGCTGGAAAGTTCACGCTGGTGGCCAGACACCTTCAACAAAGGAGCGCTCC 959
QY 161 AlaAlaSerIleAspAlaAlaIleValAlaGlyValAlaAsnGlyAsnProLeuAspAla 180
Db 960 GCGGCTTCATCCACGCGGGAAGAGCCGCGCTCAACACCGCAACCCGCTCGAGCC 1019
QY 181 ValGlnGln 183
Db 1020 GTTCAGCAG 1028
RESULT 8
AA291074
ID AA291074 standard; DNA: 564 BP.
XX
AC AA291074;
XX
DT 06-JUN-2000 (first entry)
XX
DE Streptomyces avidinii mat gene.
XX
KM Plant somatic tissue degeneration; plant essential factor; depletion;
XX viability; mat gene; plant development; plant morphology; flower;
XX fruit plant; de.
XX
OS Streptomyces avidinii.
XX
PN WO200007427-A2.
XX
PD 17-FEB-2000.
XX
PF 30-JUL-1999; 99MO-IL00420.
XX
PR 03-AUG-1998; 98IL-0125632.
XX
PA (AGRI-) AGRIC RES ORG.
XX
PI Kapulnik Y, Ginzberg I;
XX
DR MPI; 2000-195402/17.
XX
DR P-PSDB; AA050513.
XX
PT Degeneration of somatic plant tissue by expression of a heterologous
XX protein, useful for controlling plant development and morphology, such
XX as decreasing the number of flowers present to increase the number of
XX fruit -
XX
PS Examples: Page 85; 91pp; English.
XX
CC The invention relates to a method of effecting degeneration of a somatic
XX plant tissue by expressing a heterologous protein capable of binding a
XX plant essential factor (PEF), in somatic plant tissue cells, where
XX heterologous protein expression causes depletion of the PEF so the plant
XX viability is maintained, while simultaneous degeneration of the somatic
XX plant tissue is effected. This sequence represents the Streptomyces
XX avidinii mat gene as an example of a heterologous gene introduced into
XX the plants. The methods can provide for the selective and optionally
XX reversible cell degeneration in somatic plant tissue. They can be used

CC For artificially controlling plant development and morphology. They can
CC be used e.g. to decrease the number of flowers in fruit producing plants
CC so as to increase the number of fruits which reach maturity.
XX
SQ Sequence 564 BP; 109 A; 213 C; 169 G; 73 T; 0 other;
Alignment Scores:
Pred. No.: 1 42e-73 Length: 564
Score: 931.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.47% Indels: 0
DB: 21 Gaps: 0
US-09-589-870B-2 (1-183) x AA291074 (1-564)
QY 2 ArgIleValAlaValAlaIleAlaValIleSerLeuThrThrValSerIleThrAlaSer 21
Db 14 CCGAAGATGCTGCTGGACCATCGCGCTTCCCTGACCAACGGGTCCGTTACGGCCAG 73
QY 22 AlaSerAlaAspProSerIleAspSerIleValGlnValSerAlaAlaGluAlaGly 41
Db 74 GCTTCGCGACACCCCTCCAAAGACTCGAAGGCCCAAGTCTCGGCGCCGAGGCGGCATC 133
QY 42 ThrGlyThrTrpTrpAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAspGly 61
Db 134 ACCGGACCTGTGACCAACCGCTCGCTGACCTTCATGTGACCGCGGCGCGCGAGC 193
QY 62 AlaLeuThrGlyThrTyrcGluSerAlaValGlyAsnAlaGluSerArgTyValLeuThr 81
Db 194 GCCCTGACCGGAACCTAGAGTGGCGCGTGGCAACCGGAGACCGGCTACGCTCGACC 253
QY 82 GlyArgTrpAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTTPThrVal 101
Db 254 GGTCTGTACAGACGCGCCCGCCGACCGAGCGGACCGCGCCCTCGTTGGAGCGTG 313
QY 102 AlaTrpIleAsnAntyTrpAsnAlaHisSerAlaThrThrTrpSerGlyIntYrVal 121
Db 314 GCTTGAAATPACTACCGCAACCGCACCTCGGACCAACGAGCGGCGGCGGAGTACGTC 373
QY 122 GlyIleAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrGluAla 141
Db 374 GGGGGGCGGAGCGGAGGATCAACACCCAGTGGCTGTGACTCCGCGACCAACCGAGCC 433
QY 142 AsnAlaTrpIleSerThrLeuValGlyHisAspThrPheThrIleValIleProSerAla 161
Db 434 AACGCTGGAAATGCCAGCTGCTCGCGCCAGCACCTTCAACCAAGTGAAGCGCTCGCC 493
QY 162 AlaSerIleAspAlaAlaIleValIleValAlaGlyValAlaAsnGlyAsnProLeuAspAlaVal 181
Db 494 GCTTCATGACGCGCGGAGAAAGCGCGGTCAACACGGAACCGGCTCGACGCGGTT 553
QY 182 GlnGln 183
Db 554 CAGCAG 559
RESULT 9
AAQ53412
ID AAQ53412 standard; DNA: 552 BP.
XX
AC AAQ53412;
XX
DT 25-MAR-2003 (updated)
XX
DT 27-JUN-1994 (first entry)
XX
DE Streptavidin gene.
XX
XX Streptavidin; protein secretion; Bacillus subtilis; ss.
XX
KM Streptomyces avidinii.
XX
OS Streptomyces avidinii.
XX
FH Key Location/Qualifiers
FT CDS 1..552

```

FT      mat_peptide      73..519      /*tag= a
FT      sig_peptide      1..72      /*tag= b
FT      misc_feature      109..519  /*tag= c
FT      /note= "fused to npr signal peptide in pBE673"
FT      W09324631-A1.
XX
XX      09-DEC-1993.
XX
XX      27-MAY-1993;      93WO-US05240.
XX
XX      29-MAY-1992;      92US-0891524.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
XX      NagaraJan V;
XX
XX      WPI, 1993-405822/50.
XX      P-PSDB; AAR44491.
XX
XX      Streptavidin prodn. from Bacillus subtilis - using signal protein
XX      from bacterial exo-protein and expression element from Gram
XX      positive bacterial protein.
XX
XX      Disclosure; Fig 1b; 54pp; English.
XX
XX      CC      Tetrameric biologically active streptavidin is produced by secretion
XX      from Bacillus subtilis transformed with a plasmid encoding the
XX      CC      sequence.
XX      CC      (updated on 25-MAR-2003 to correct PN field.)
XX
XX      SQ      Sequence 552 BP; 105 A; 207 C; 167 G; 70 T; 3 other;
XX
XX      Alignment Scores:
XX      Pred. No.:      3,12e-73      Length:      552
XX      Score:      927.00      Matches:      182
XX      Percent Similarity:      99.45%      Conservative:      0
XX      Best Local Similarity:      99.45%      Mismatches:      1
XX      Query Match:      99.04%      Indels:      0
XX      DB:      14      Gaps:      0
XX
US-09-589-870b-2 (1-183) x AAG53412 (1-552)
QY      1 MetArgLysIleValValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
DB      1 ATGCGCAAGATCGTCGTGCGATCGCGCTTCCCTGACACAGGTCGATTACGGCC 60
QY      21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGlnAgl 40
DB      61 AGGCGCTTCGCGAGATGCTCCAGAGACTCGAAGCGCCGAGTCTCGCGCCGACGCGCC 120
QY      41 IleThrGlyThrTyrPzAsnGlnLeuGlySerThrPheIleValThrAlaGlnValAsp 60
DB      121 ATCACCGGACCTGATGACACAGCTCGGCTGACCTTCATGATGACGCGGCGCCGAC 180
QY      61 GlyAlaLeuThrGlyThrTyrGlnSerAlaValGlyAsnAlaGlnSerArgTyrValLeu 80
DB      181 GCGCGCCCTGACCGGAACTACGAGTCGCGCGGCAACGCCGACGCGCTGCTGCTG 240
QY      81 ThrGlyArgTyrAsnSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
DB      241 ACCGGTCGTTACGACAGCGCGCCCGGACCGACGCGGACCGGACCGCTCGGTTGGAG 300
QY      101 ValAlaTyrPzAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGlyGlnTyr 120
DB      301 GTGGCTCGAAGATATACCTACCGCAACGCCCACTCGGAGCACCTGAGAGCGCGCAGTAC 360
QY      121 ValGlyGlyAlaGlnAlaArgIleAsnThrGlnTyrPleuLeuThrSerGlyThrThrGln 140

```

```

DB      361 GTCGCGCGCGCGGAGCGAGATCAACACCACTGCTGTGACTCTCCGCGACACCGAG 420
QY      141 AlaAsnAlaTyrPzSerThrIleuValGlyHsAspThrPheThrLysValLysProSer 160
DB      421 GCGCAAGCTTGAGAGTCCAGCTGTGCGGCGACGACCTTCACCAAGGTGAAGCGCTCC 480
QY      161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAlaAsnGlyAsnProLeuAspAla 180
DB      481 GCGGCTTCATGACCGCGCGGAGAGAGCGCGCTCAACACGCGCAACCGCTCGACGCC 540
QY      181 ValGlnGln 183
DB      541 GTTCAGCAG 549
RESULT 10
AAV34715
ID      AAV34715 standard; DNA; 638 BP.
AC      AAV34715;
XX
XX      27-AUG-1998 (first entry)
XX
XX      S. avidinii streptavidin mutant #1 DNA.
XX
XX      Streptavidin; ligand; binding affinity; mutant; isolation;
XX      purification; recover; immobilise; ss.
XX
XX      Streptomyces avidinii.
XX      Synthetic.
XX
XX      Key      Location/Qualifiers
XX      CDS      50..601
XX      sig_peptide      /*tag= a
XX      mat_peptide      122..598      /*tag= b
XX      /*tag= c
XX      /product= Streptavidin
XX      mutation      251..262
XX      /*tag= d
XX      /note= "wild type GAGTCGCGCGTC is replaced by
XX      GTCACGCGCGCT"
XX
EP835934-A2.
XX
XX      15-APR-1998.
XX
XX      09-OCT-1997;      97EP-0117504.
XX
XX      10-OCT-1996;      96DE-1041876.
XX
XX      (BIOA-) INST BIOANALYTIK GMBH GOETTINGEN.
XX
XX      Skerra A, Voss S;
XX      WPI, 1998-218868/20.
XX      P-PSDB; AAW59217.
XX
XX      Streptavidin mutants with higher binding affinity for peptide
XX      ligands - have mutation in amino acid region 44-53, used to isolate,
XX      purify or determine fusion proteins including these ligands
XX
XX      Disclosure; Page -; 21pp; German.
XX
XX      This sequence encodes a mutant streptavidin protein isolated from
XX      Streptomyces avidinii where the residues E54V at position 44-47 of
XX      the mature wild type sequence are replaced by YTRK. This sequence is
XX      used to produce mutants which are used in a method to assay the binding
XX      affinity of streptavidin mutants. These mutants have a mutation within
XX      the amino acid (aa) region 44-53 of the wild-type protein show a higher
XX      binding affinity than the wild-type for peptide ligands that include
XX      the sequence of formula Trp-X-His-Pro-Gln-Phe-Y-Z where X = any aa; Y
XX      and Z are both Gly, or Y = Gln and Z = Arg or Lys. Recombinant

```

streptavidin mutants can be used to isolate, purify and determine proteins or to determine/recover substances that contain streptavidin-binding groups. Such compounds may also be used to immobilise fusions on microtitre plates, microbeads or sensor chips. NOTE: This sequence does not appear in the specification but has been constructed from the wild-type streptavidin sequence represented in AAV34714.

SO Sequence 638 BP; 115 A; 244 C; 193 G; 86 T; 0 other;

Alignment Scores:

Pred. No.:	1,88e-72	Length:	638
Score:	919.00	Matches:	180
Percent Similarity:	98.91%	Conservative:	1
Best Local Similarity:	98.36%	Mismatches:	2
Query Match:	98.18%	Indels:	0
DB:	19	Gaps:	0

US-09-589-870B-2 (1-183) x AAV34715 (1-638)

```
QY 1 MeArGLyRIleValAlaAlaIleAlaValSerLeuThrThrValSerIleThraAla 20
    |||
DB 50 ATGCGCAAGATCGTCTGACGACATCGCGTTCCCTGACCGGCTTCGATTACGGCC 109

QY 21 SerAlaSerAlaAaPProSerLySaPSeLyAlaGlnValSerAlaAlaGlnAlaGly 40
    |||
DB 110 ACGGCTTCGGCAGACCCCTCCAGACTCGAAGGCCAGCTCCGGCCCGCGAGCCGGC 169

QY 41 IleThrGlyThrTrpTyraGlnIleuGlySerThrPheIleValThrAlaGlyAlaAaP 60
    |||
DB 170 ATCACCAGGACCTGATACAAACAGCTCGGCTCAGCTTCTGACCGGCGGCGCGAC 229

QY 61 GlyAlaIleuThrGlyThrTyrgIuSerAlaValAlaGlnAlaIleuSerArgTyraIleu 80
    |||
DB 230 GGGCGCCCTGACCGAAGCTAAGTACGCGCGCGCGCGCGCGCGCGCTACGCTCTG 289

QY 81 ThrGlyArgTyraPseSerAlaProAlaThrAaPGLySerGlyThrAlaIleuGlyTrpThr 100
    |||
DB 290 ACCGCTCGTTACACAGCGCCCGCGCGCGCGCGCGCGCGCGCGCTTCGATTGAGC 349

QY 101 ValAlaTrpLySaPSeLySaPSeLySaPSeLySaPSeLySaPSeLySaPSeLySaPSe 120
    |||
DB 350 GTGGCTGGAAAGTAATACACCGCAACGCCCACTCCGCGCGCGCGCGCGCGCGCGCTAC 409

QY 121 ValGlyValAlaGlnAlaAaGlnIleuThrGlnIleuThrGlnIleuThrGlnIleu 140
    |||
DB 410 GTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 469

QY 141 AlaAaAlaTrpLySaPSeLySaPSeLySaPSeLySaPSeLySaPSeLySaPSeLySaPSe 160
    |||
DB 470 GCCAAGCGCTGGAAAGTACCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 529

QY 161 AlaAlaSerIleAaPAlaAlaValLySaPSeLySaPSeLySaPSeLySaPSeLySaPSe 180
    |||
DB 530 GCCGCTTCATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 589

QY 181 ValGlnGln 183
    |||
DB 590 GTTCACAGCAG 598
```

RESULT 11
AAV34716
ID AAV34716 standard; DNA; 638 BP.

XX AAV34716;
XX
DT 27-AUG-1998 (first entry)
XX
DE S. avidinii streptavidin mutant #2 DNA.
XX
KM Streptavidin; ligand; binding affinity; mutant; isolation;
XX purification; recover; immobilise; ss.

OS Streptomyces avidinii.
OS Synthetic.

Key	Location/Qualifiers
Key	50..601
CDS	/tag= a
sig_peptide	50..121
mat_peptide	/tag= b
mat_peptide	122..598
FT	/tag= c
FT	/product= Streptavidin
FT	251..262
FT	/tag= d
FT	/note= "Wild type GAGTCGCGCGCGC is replaced by ATCGTCGAGC"

EP835934-A2.

15-APR-1998.

09-OCT-1997; 97EP-0117504.

10-OCT-1996; 96DE-1041876.

(BIOA-) INST BIOANALYTIK GMBH GOETTINGEN.

Skerra A, Voss S;

WPI: 1998-218868/20.

P-PSDB; AAW59218.

Streptavidin mutants with higher binding affinity for peptide ligands - have mutation in amino acid region 44-53, used to isolate, purify or determine fusion proteins including these ligands

Disclosure; Page -: 21pp; German.

This sequence encodes a mutant streptavidin protein isolated from Streptomyces avidinii where the residues E54V at position 44-47 of the mature wild type sequence are replaced by IGAR. This sequence is used to produce mutants which are used in a method to assay the binding affinity of streptavidin mutants. These mutants have a mutation within the amino acid (aa) region 44-53 of the wild-type protein show a higher binding affinity than the wild-type for peptide ligands that include the sequence of formula Trp-X-His-Pro-Gln-Phe-Y-Z where X = any aa; Y and Z are both Gly, or Y = Glu and Z = Arg or Lys. Recombinant streptavidin mutants can be used to isolate, purify and determine proteins or to determine/recover substances that contain streptavidin-binding groups. Such compounds may also be used to immobilise fusions on microtitre plates, microbeads or sensor chips. NOTE: This sequence does not appear in the specification but has been constructed from the wild-type streptavidin sequence represented in AAV34714.

SO Sequence 638 BP; 116 A; 242 C; 194 G; 86 T; 0 other;

Alignment Scores:	2,83e-72	Length:	638
Pred. No.:	917.00	Matches:	180
Score:	917.00	Conservative:	0
Percent Similarity:	98.36%	Mismatches:	3
Best Local Similarity:	97.97%	Indels:	0
Query Match:	97.97%	Gaps:	0
DB:	19		

US-09-589-870B-2 (1-183) x AAV34716 (1-638)

```
QY 1 MeArGLyRIleValAlaAlaIleAlaValSerLeuThrThrValSerIleThraAla 20
    |||
DB 50 ATGCGCAAGATCGTCTGACGACATCGCGTTCCCTGACCGGCTTCGATTACGGCC 109

QY 21 SerAlaSerAlaAaPProSerLySaPSeLySaPSeLySaPSeLySaPSeLySaPSeLySaPSe 40
    |||
DB 110 ACGGCTTCGGCAGACCCCTCCAGACTCGAAGGCCAGCTCCGGCCCGCGAGCCGGC 169
```

Qy	41	IllethrGlyThrTrpTyrAsnGlnLeuGlySerThrPheIleValThrAlaGlyValAsp	60
Db	170	ATCCAGGACCTGTGTACCAACCAAGCTGGCTCGACTTCATCTGACCGCGGGCGGAC	222
Qy	61	GlyAlaLeuThrGlyThrTyrGlnSerAlaValGlyAsnAlaGlnSerArgTyrValLeu	80
Db	230	GGCGCCTGAACCGGAACTCATGTCGGTCCAGGGGCAACCGCGAGCCCTTACGCTTG	288
Qy	81	ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr	100
Db	290	ACCGTCGTATACGACACAGCGCCCGCCACCGAGCGGACCGCCCTCGTTGAGACG	349
Qy	101	ValAlaTrpLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyr	120
Db	350	GTCGGCTCGAGAGATACTAACCGCAAGCCCACTCCGCGAACAGTGAGCGGCAAGTAC	409
Qy	121	ValGlyGlyValaGlnAlaArgIleAsnThrGlnThrLeuLeuThrSerGlyTyrThrGlu	140
Db	410	GTCGGCGGCGCCAGGAGGAGATCAACACCAGTGGTGTGACTTCGGCACACCGAG	469
Qy	141	AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrIlysValLysProSer	160
Db	470	GCCAAAGCCCTGGAAGTCCAGCTGTGCGCACAGACCTTTCACCAAGGTGAAGCGTCC	529
Qy	161	AlaAlaSerIleAspAlaAlaIleLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla	180
Db	530	GCGGCTTCATCGACCGCGGAGAAAGCGCGCTCAACAAAGGCAACCGCTCGAGGCC	589
Qy	181	ValGlnGln 183	
Db	590	GTTTCAGCAG 598	
RESULT 12			
ID	AAC86563 standard; DNA; 1239 BP.		
XX	AAC86563;		
AC	AAC86563;		
DT	02-APR-2001 (first entry)		
XX			
DE	DNA encoding a fusion of anti-CD20 single chain antibody/streptavidin.		
XX			
KW	Streptavidin; tumour cell; cancer; adenocarcinoma;		
KW	hematological malignancy; ss.		
XX			
OS	Synthetic.		
OS	Streptomyces avidinii.		
OS	Homo sapiens.		
XX			
PN	WO200075333-A1.		
XX			
PD	14-DEC-2000.		
XX			
PE	05-JUN-2000; 2000WO-US15595.		
XX			
FR	07-JUN-1999; 99US-0137900.		
PR	03-DEC-1999; 99US-0168976.		
XX			
PA	(NEOR-) NEORX CORP.		
XX			
PI	Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;		
DR	WPI; 2001-091213/10.		
DR	P-PSDB; AAB30694.		
XX			
PT	New vector constructs for expressing genomic streptavidin fusion		
XX	proteins which are useful for targeting tumour cells associated with		
XX	cancer. e.g. adenocarcinomas -		
XX			
PS	Example 2; Fig 11A; 100pp; English.		
XX			
CC	The present sequence encodes a fusion of an anti-CD20 single chain		

antibody (B99) streptavidin. The fusion protein is expressed using vectors of the invention. The specification describes vector constructs for expressing streptavidin fusion proteins. The vector comprises a nucleic acid encoding streptavidin or its functional variant operatively linked to a promoter, and a cloning site for insertion of a second nucleic acid sequence encoding a polypeptide to be fused with streptavidin. Interposed between the promoter and the first nucleic acid sequence. Alternatively, the vector construct comprises a nucleic acid, operatively linked to a promoter, encoding a polypeptide to be fused with streptavidin, and a cloning site for insertion of a second nucleic acid encoding at least 129 amino acids of streptavidin or its functional variant. The fusion proteins are useful for targeting tumour cells, particularly tumour cells associated with cancer. e.g. adenocarcinomas or hematological malignancies. The vector construct is useful for expressing of streptavidin fusion proteins. In particular, these are useful as tools for medical diagnostics and therapeutic purposes, e.g. for detecting the presence or absence of, or treating, a target site within a mammalian host.

Sequence 1239 BP; 270 A; 392 C; 356 G; 221 T; 0 other;

Alignment Scores:	
Pred. No.:	1,686-66
Score:	85.50
Percent Similarity:	98.83%
Best Local Similarity:	99.48%
Query Match:	91.40%
DB:	22
Gaps:	1
Length:	1239
Matches:	165
Conservative:	4
Mismatches:	1
Indels:	1
Gaps:	1

US-09-589-870B-2 (1-183) x AAC86563 (1-1239)

QY	14	ThrThtValSerIleThrIaL---SerIaSerIaIaAspProSerIyAspSerIyAla	32
Dd	724	ACCACGGGTACCGGTAGCTCTGCTGTTGGCGACACCCCTCCAGAGACTCGAAGGCGC	783
QY	33	GlnValSerIaIaIaGlnValaGlyIleThrGlyThrTPTyAsnGlnIleuGlySerThr	52
Dd	784	CAGGTCTCTGGCGCCGAGGCGCGGCAATCACCGGCACCTGGTACACACAGCTCGGCTCGACC	843
QY	53	PheIleValIleThrAlaGlyValaAspGlyIaIleuThrGlyThrTtyGluSerAlaValGly	72
Dd	844	TTTCATCGTACCGCGGGGCGCGACGCGCCCTGACCGGAACCTTAGAGTCGGCCGTCCGGC	903
QY	73	AsnAlaGluSerArgTyrValIleuThrGlyArgTyrAspSerAlaProAlaThrAspGly	92
Dd	904	AACGCCGAGAGCGCGTACGTCCTGACCGGTCGTTACGACAGCGGCCCGGCCACCGACGGC	963
QY	93	SerGlyThrAlaIleuGlyTTPThrThyValaIaIaTryptIyAsnAsnTyrArgAsnAlaHisSer	112
Dd	964	AGCGGACCGCCCTCGGTGGACGGTGGCCCTGGAAAGTAATCATCCGCAACGCCACATCC	1022
QY	113	AlaIaThrThrTTPSerGlyGlnTyrValaGlyGlyAlaGlnAlaArgIleAsnThrGlnTrp	132
Dd	1024	GCGACCACTGTGAGAGGGCCAGTAGCTGGCGGGCCGCGACGAGATCAACCCACCATGG	1083
QY	133	LeuIleuThrSerGlyThrThrGlnAlaAsnAlaIaTryptIySerThrLeuValGlyHisAsp	152
Dd	1084	CTGCTGACCTCCGACACACCGAAGCCCAACCCCTGGAAAGTCCAGCGTGGTGGCGCACGAC	1144
QY	153	ThrPheThrIyValIyProSerAlaIaSerIleAspIaIaIaIaIyLysAlaGlyAla	172
Dd	1144	ACCTTCACCAAGGTGAAGCCGCTCCGCCGCCCTTCATTCGACGCGCGGAGAGACCGGCGCTC	1203
QY	173	AsnAsnGlyAsnProLeuAspAlaValaGlnGln	183
Dd	1204	AACAAACGCAACCCGCTCGAAGCGCGTTCAGCAG	1236
RESULT 13			
AAC86562			
ID AAC86562 standard; DNA; 1612 BP.			
XX AAC86562;			
AC			

```

01 02-APR-2001 (first entry)
02 DNA encoding a fusion of single chain antibody/streptavidin.
03
04 Streptavidin; tumour cell; cancer; adenocarcinoma;
05 hematological malignancy; huNR-LU-10; EGP40; EPCAM; ss.
06
07 Synthetic.
08 Streptomyces avidinii.
09
10 Homo sapiens.
11
12 Key location/Qualifiers
13 CDS 311..1606
14 FT /tag= a
15 FT /*product= "huNR-LU-10 and streptavidin fusion"
16
17 MO200075333-A1.
18
19 14-DEC-2000.
20
21 05-JUN-2000; 2000MO-US15595.
22
23 07-JUN-1999; 99US-0137900.
24 03-DEC-1999; 99US-0168976.
25
26 (NEOR-) NEORX CORP.
27
28 Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JW,
29 WPI: 2001-091213/10.
30 P-PSDB; AAB30693.
31
32 New vector constructs for expressing genomic streptavidin fusion
33 proteins which are useful for targeting tumour cells associated with
34 cancer, e.g. adenocarcinomas -
35
36 Example 1; Fig 10; 100pp; English.
37
38 The present sequence encodes a fusion of the single chain antibody
39 huNR-LU-10 and streptavidin. The antibody binds the antigen EGP40 or
40 EPCAM. The fusion protein is expressed using vectors of the invention.
41
42 The specification describes vector constructs for expressing streptavidin
43 fusion proteins. The vector comprises a first nucleic acid encoding
44 streptavidin or its functional variant operatively linked to a promoter,
45 and a cloning site for insertion of a second nucleic acid sequence
46 encoding a polypeptide to be fused with streptavidin, intersposed between
47 the promoter and the first nucleic acid sequence. Alternatively, the
48 vector construct comprises a first nucleic acid, operatively linked to a
49 promoter, encoding a polypeptide to be fused with streptavidin, and a
50 cloning site for insertion of a second nucleic acid encoding at least
51 129 amino acids of streptavidin or its functional variant. The fusion
52 proteins are useful for targeting tumour cells, particularly tumour cells
53 associated with cancer, e.g. adenocarcinomas or hematological
54 malignancies. The vector construct is useful for expressing of
55 streptavidin fusion proteins. In particular, these are useful as tools
56 for medical diagnosis and therapeutic purposes, e.g. for detecting the
57 presence or absence of, or treating, a target site within a mammalian
58 host.
59
60 Sequence 1612 BP; 353 A; 493 C; 459 G; 307 T; 0 other;
61
62 Alignment Scores:
63 Pred. No.: 2,8e-66 Length: 1612
64 Score: 854.50 Matches: 166
65 Percent Similarity: 97.67% Conservative: 2
66 Best Local Similarity: 96.51% Mismatches: 3
67 Query Match: 91.29% Indels: 1
68 Gaps: 1
69
70 US-09-569-870B-2 (1-183) X AACG6562 (1-1612)

```

Db	1091	ACCTTACTCACCGGAGC---TCTGGCTCTGGTTTGGCGACAACCCCTTCCAAAGACTCGAAG	1147
Oy	32	AlaGlnValSerAlaAlaGlnAlaGlyIleThrGlyThrTrpTyraGlnIleuGlySer	51
Db	1148	GCCCAAGGCTTCGGCGCGCGAGGCCGGGATCATCCGGGACCTGGTAAACAACGAGCTCGGCTCG	1207
Oy	52	ThrIleIleValAlThrAlaGlyAlaAPGlyValAlaIleuThrGlyThrTyrgIuSerAlaVal	71
Db	1208	ACCTTCATCGGAGCCGGGGCGCGGACGGGGCCCTTGACCCGGAACCTACGAGTCGGCCGTC	1267
Oy	72	GlyValnAlaGluSerArgTyrrValIleuThrGlyArgTyrrAspSerAlaProAlaThrAsp	91
Db	1268	GCGAACCGCGAGACCCCTCACTGCTCTGACCGGATCTTACGACAGCGCCCGCGCACCGAC	1327
Oy	92	GlySerGlyThrAlaIleuGlyTrrThrValAlaTrpIysAsnAsnTyrrAlaGlnAlaHis	111
Db	1328	GGCAGCGGCGACCGCGCTCGGTTGGACGGTGGCTCGCTGAAAGATTAACATACCGCAACGCCAC	1387
Oy	112	SerAlaThrThrTrpSerGlyGlnTyrrValGlyGlyAlaGlnAlaArgIleAsnThrGln	131
Db	1388	TCCCCGACCACTGTGACCGGCGCGAGTACGTGCGGGCGCGCGAGCGAGATCAACACCCAG	1447
Oy	132	TrpIleuThrTrpSerGlyThrThrGlnAlaAsnAlaTrpIysSerThrIleuValGlyHis	151
Db	1448	TGGCTGCTGACCTTCGGGACACCGAGGCCAACGCCCTGGAAATCCACGCTGTGTGGCAC	1507
Oy	152	AspThrPheThrIysValIlySPProSerAlaAlaSerIleAspAlaAlaIlySLysAlaGly	171
Db	1508	GACACCTTCACCAAGGTGAAGCCGTCGCGCGCTCCATCCACCGCGGGAAGAGCGCGGC	1567
Oy	172	ValAsnAsnGlyAsnProIleuAspAlaValGlnGln	183
Db	1568	GTCAAACAACGGCAACCCGCTCGACGCGCTTCACAGAC	1603
RESULT 14			
AAc86564	AAC86564 standard; DNA; 1280 BP.		
XX	AC	AAC86564;	
XX	DT	02-APR-2001 (first entry)	
XX	DE	Anti-CD20 single chain antibody/streptavidin fusion protein cassette.	
XX	KW	Streptavidin; tumour cell; cancer; adenocarcinoma;	
XX	KW	hematological malignancy; ss.	
XX	OS	Synthetic.	
XX	OS	Streptomyces avidinii.	
XX	OS	Homo sapiens.	
FT	Key	Location/Qualifiers	
FT	CDS	3..1274	
FT		/*tag= a	
FT		/product= "anti-CD20 scFv and streptavidin fusion"	
XX	PM	MO200075333-A1.	
XX	PD	14-DEC-2000.	
XX	PF	05-JUN-2000; 2000MO-US15595.	
XX	PR	07-JUN-1999; 99US-0137900.	
XX	PR	03-DEC-1999; 99US-0168976.	
XX	PA	(NEOR-) NEORX CORP.	
XX	PI	Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;	
XX	DR	WPI: 2001-091213/10.	
XX	DR	P-PSDB; AAB30695.	
XX	PT	New vector constructs for expressing genomic streptavidin fusion	

PT proteins which are useful for targeting tumour cells associated with
 cancer, e.g. adenocarcinomas
 XX
 PS
 XX Example 2; Fig 11C; 100bp; English.

CC The present sequence encodes a fusion of an anti-CD20 single chain
 CC antibody (B99) streptavidin. The fusion protein is expressed using
 CC vectors of the invention. The specification describes vector constructs
 CC for expressing streptavidin fusion proteins. The vector comprises a
 CC nucleic acid encoding streptavidin or its functional variant operatively
 CC linked to a promoter, and a cloning site for insertion of a second
 CC nucleic acid sequence encoding a polypeptide to be fused with
 CC streptavidin, interposed between the promoter and the first nucleic
 CC acid sequence. Alternatively, the vector construct comprises a nucleic
 CC acid, operatively linked to a promoter, encoding a polypeptide to be
 CC fused with streptavidin, and a cloning site for insertion of a second
 CC nucleic acid encoding at least 129 amino acids of streptavidin or its
 CC functional variant. The fusion proteins are useful for targeting tumour
 CC cells, particularly tumour cells associated with cancer.
 CC e.g. adenocarcinomas or hematological malignancies. The vector construct
 CC is useful for expressing of streptavidin fusion proteins. In particular,
 CC these are useful as tools for medical diagnostics and therapeutic
 CC purposes, e.g. for detecting the presence or absence of, or treating, a
 CC target site within a mammalian host.

SO Sequence 1280 BP; 267 A; 397 C; 388 G; 228 T; 0 other;

Alignment Scores:

Pred. No.:	9,81e-66	Length:	1280
Score:	847.00	Matches:	162
Percent Similarity:	98.79%	Conservative:	1
Best Local Similarity:	98.18%	Mismatches:	2
Query Match:	90.49%	Indels:	0
DB:	22	Gaps:	0

US-09-589-870B-2 (1-183) x AAC86564 (1-1280)

QY 19 ThAlaSerAlaSerAlaSPProSerLyASpSerLyAlaGlnValSerAlaAlaGln 38
 DB 777 TCTGGCTCTGTTCTGGCAGACCCCTCAAGACTCCAGAGGCTCTGGCCGCGAG 836
 QY 39 AlaGlyIleThrGlyThrTrpTyraSngInLeuGlySerThPheIleValhrAlaGly 58
 DB 837 GCGGATACCGGACCTGTTACCAACAGCTGCTGCTTCACTTCATCTGACCCGCGGCG 896
 QY 59 AlaSPGlyAlaLeuThrGlyThrTyrgInuSerAlaValGlyASnaIaGlnuSerArgTy 78
 DB 897 GCCGAGCGCCCTGACCGAAGCTAGAGTCCGCGCTGCGAAGCGGAGCGGCTAC 956
 QY 79 ValIleThrGlyArgTyraSPSerAlaProAlaThrAspGlySerGlyThrAlaLeuGly 98
 DB 957 GTCTGACCGGTGTTACAGACGCGCCCGCCACCGAGGAGGCGGACCCCTCGGT 1016
 QY 99 TrpThrValAlaTrpLyAsnaSntTyraGnaAlaAhISeSerAlaThThrTrpSergly 118
 DB 1017 TGGACGCTGCTGGAGAAATACTACCGCAAGCCCACTCCGCAACACGTGGAGCGGC 1076
 QY 119 GlnTyraValGlyAlaGlnAlaArgIleAsnThrgInTrpLeuThrSerglyThr 138
 DB 1077 CAGTACGTGGCGGCGCCGAGCGAGATCAACACCCAGTGGCTCTGACCTCCGCGACC 1136
 QY 139 ThrGlnAlaAsnaIaTrpLySerThrLeuValGlyHISAspThrPheThrLyValys 158
 DB 1137 ACCGAGGCCAACGCTGGAAGTCCAGCTGCTGGCCACGACACCTTACCAAGGTGAG 1196
 QY 159 ProSerAlaIaSerIleAspAlaIaLySlyAlaGlyValaAsnAGIYAsnProLeu 178
 DB 1197 CGGTCCGCGCTCATCGACCGCGGAGAAAGCGCGCTCAACACGGAACCCGCTC 1256
 QY 179 AspAlaValGlnIn 183
 DB 1257 GACGCGCTTACAGAC 1271

RESULT 15
 AAQ70659
 ID AAQ70659 standard; DNA; 1356 BP.

XX
 AC AAQ70659;
 XX
 DT 25-MAR-2003 (updated)
 DT 26-MAR-1995 (first entry)
 XX
 DE ScfV pRAS108 and pRAS112 DNA.

XX Amplification; single chain variable region fusion protein; PCR; ss.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT CDS 40..1348
 FT /*tag= a

PN WO9415644-A1.

XX 21-JUL-1994.

PF 17-JAN-1994; 94WO-GB00087.

PR 15-JAN-1993; 93GB-0000686.

PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.

XX Deonaraia M, Epenetos AA, Spooner RA;

XX WPI, 1994-248907/30.

DR P-PSDB; AAR56483.

XX New cpds. comprising a targeting portion and a cytotoxic portion

PT - used esp. for treating mammals for destroying target cells,

PT partic. tumour cells

XX Claim 36; Fig 9; 114pp; English.

CC The sequence is that of the ScfV pRAS108 and pRAS112 between HindIII and

CC EcoRI site obtd. by PCR.

CC See also AAQ70652-69.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1356 BP; 332 A; 359 C; 359 G; 306 T; 0 other;

Alignment Scores:

Pred. No.:	1.29e-65	Length:	1356
Score:	846.00	Matches:	162
Percent Similarity:	99.39%	Conservative:	1
Best Local Similarity:	98.78%	Mismatches:	1
Query Match:	90.38%	Indels:	0
DB:	15	Gaps:	0

US-09-589-870B-2 (1-183) x AAQ70659 (1-1356)

QY 20 AlaSerAlaSerAlaSPProSerLyASpSerLyAlaGlnValSerAlaAlaGln 39
 DB 853 GCTGCCACCTGACAGACCCGTCMAAGACTCCAAAGCTTAGGTTTCTGAGCGGAAGCT 912
 QY 40 GlyIleThrGlyThrTrpTyraSngInLeuGlySerThPheIleValThralaGlyAla 59
 DB 913 GGTATCACTGGACCTGCTGTTAACAACCTGGGTGCTTATTTGACCCGCTGGTGGC 972
 QY 60 AspGlyAlaLeuThrGlyThrTyrgInuSerAlaValGlyASnaIaGlnuSerArgTy 79
 DB 973 GACGAGCTTGTGCTGACCTAGATCGCGTTGGTAAACCAAAATCCCTTACGTA 1032
 QY 80 LeuThrGlyArgTyraSPSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrp 99
 DB 1033 CTGACTGGCGGCTTATGACTGTGACCTGCCAGGATGGCTTGGTACCGCTGGGCTGG 1092

```
QY 100 ThrValAlaTrpLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGlyGln 119
Db 1093 ACTGTGGCTTGGAAAAACAACATCGTAATGCCACAGCCCACTAGTGTCTGGCCAA 1152
QY 120 TyrValGlyGlyValAlaGlyAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThr 139
Db 1153 TACGTTGGCGGCTGCTGAGCGCTGATCAACACTCAGTGGCTGTTAACATCGGCACTACC 1212
QY 140 GluAlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysPro 159
Db 1213 GAAGCGAATGATGAAATGACACTAGTAGGTCAATGACACCTTTACCAAGTTAAGCCT 1272
QY 160 SerAlaAlaSerIleAspAlaAlaLysValAlaGlyValAsnAsnGlyAsnProLeuAsp 179
Db 1273 TCTGCTGCTAGCATTTGATGCTGCCAAGAAAGCAGCGCTAAACACGCTAACCTCTAGAC 1332
QY 180 AlaValGlnGln 183
Db 1333 GCTGTTACGCAA 1344
```

Search completed: October 28, 2003, 02:09:54
Job time : 257 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 27, 2003, 21:54:43 ; Search time 3292 Seconds

(without alignments)
2274.138 Million cell updates/sec

Title: US-09-589-870B-2

Sequence: 1 MKKIVAAIAVSLTVSIFA.....IDAKKAGVGNNGPLDAVQ 183

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=klh
-Q=/cgn2.1/USPTO.epoc1/US09589870/runac_27102003.104445.16617/app_query.fasta.1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09589870@CNC1_1_3508@runac_27102003.104445.16617 -NCPU=6 -ICPU=3
-NO_WMAP -LARGECUERY -NEG SCORES=0 -WAIT -OSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	936	100.0	638	1	SASTRAVG	X03591 Streptomyces A91649 Sequence 1
2	936	100.0	638	6	AR177252	AR177252 Sequence 1
3	936	100.0	638	6	AR204862	AR204862 Sequence 1
4	936	100.0	638	6	AR217936	AR217936 Sequence 1
5	936	100.0	638	6	AX057941	AX057941 Sequence 1
6	936	100.0	638	6	AX057941	AX057941 Sequence 2
7	933	99.7	638	6	101349	101349 Sequence 2
8	932	99.6	625	1	S78777	S78777 streptavidin
9	897	95.8	625	1	S78782	S78782 streptavidin
10	855.5	91.4	1239	6	AX057945	AX057945 Sequence 1
11	854.5	91.3	1612	6	AX057943	AX057943 Sequence 1
12	847	90.5	1280	6	AX057947	AX057947 Sequence 1
13	846	90.4	1356	6	A39565	A39565 Sequence 1
14	846	90.4	1356	6	A93152	A93152 Sequence 4
15	846	90.4	1356	6	AR082490	AR082490 Sequence 1
16	837.5	88.5	525	6	A20698	A20698 Fxa-StrpA D
17	837.5	88.5	525	6	I15647	I15647 Sequence 2
18	835	89.2	498	6	I09059	I09059 Sequence 38
19	834	89.1	483	12	AF283893	AF283893 Synthetic
20	834	89.1	498	6	A93838	A93838 Sequence 7
21	834	89.1	507	6	A00743	A00743 S.avidinii
22	834	89.1	507	6	A00744	A00744 S.avidinii
23	834	89.1	1266	6	AX473375	AX473375 Sequence 1
24	772	82.5	1881	6	BD144212	BD144212 Fused pro
25	745	79.6	1296	6	A39567	A39567 Sequence 6
26	745	79.6	1296	6	A93154	A93154 Sequence 6
27	745	79.6	1296	6	AR082491	AR082491 Sequence 1
28	687.5	73.5	1173	6	AR231457	AR231457 Sequence 1
29	686	72.9	864	6	BD140614	BD140614 Bifunctional
30	682	72.9	1176	6	AR231456	AR231456 Sequence 1
31	682	72.9	4354	12	ASY14583	ASY14583 Artificial
32	682	72.9	4363	12	SCFV18290	SCFV18290 Expression
33	681.5	72.8	1257	6	A39569	A39569 Sequence 8
34	681.5	72.8	1257	6	A93156	A93156 Sequence 8
35	681.5	72.8	1257	6	AR082492	AR082492 Sequence 8
36	680	72.6	399	12	AGSTREP	X65082 Synthetic 9
37	679	72.5	384	6	AX348055	AX348055 Sequence 1
38	679	72.5	384	6	AX353922	AX353922 Sequence 1
39	679	72.5	1254	12	AR148718	AR148718 Synthetic
40	679	72.5	2019	6	E10169	E10169 DNA encodin
41	675	72.1	384	6	A93663	A93663 Sequence 15
42	675	72.1	384	6	AR177264	AR177264 Sequence 1
43	675	72.1	384	6	AR217948	AR217948 Sequence 1
44	675	72.1	387	6	I17472	I17472 Sequence 1
45	660	70.5	2055	6	E10170	E10170 DNA encodin

ALIGNMENTS

RESULT 1

SASTRAVG LOCUS 638 bp DNA linear BCT 30-MAR-1995
DEFINITION Streptomyces avidinii gene for streptavidin.
ACCESSION X03591
VERSION X03591.1 GI:46740
KEYWORDS streptavidin.
SOURCE Streptomyces avidinii
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Streptomycinae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 638)
Argarana, C.E., Kuntz, I.D., Birken, S., Axel, R. and Cantor, C.R.
Molecular cloning and nucleotide sequence of the streptavidin gene
Nucleic Acids Res. 14 (4), 1871-1882 (1986)
JOURNAL MEDLINE 86148514
PUBMED 3951999
COMMENT Data kindly reviewed (13-JUL-1986) by C. Argarana.
FEATURES
source location/Qualifiers
1..638
/organism="Streptomyces avidinii"
/mol_type="genomic DNA"
/db_xref="taxon:1895"
50..601
/note="put. streptavidin precursor"
/codon_start=1
/transl_table=1
/protein_id="CAA27265.1"
/db_xref="GI:46741"
/translation="MRKIVAAIAVSLTIVSTASADPSKDSKQVSAARAGITGT
WYNQGSTFIVTAGDGLTGYESAVGNAESRYVLGRYDAPATDSGTALGMWIVA
KNNYRNHSAITWGGVGAERINTQWLTSSTTEANAKSTLVGHDTFTKVKPS
AASIDAKKAGVNNGNPLDAVQ"
50..121
/note="put. signal peptide (aa -24 to -1)"
122..598
/product="mat. streptavidin (aa 1-160)"
BASE COUNT 115 a 244 c 193 g 86 t
ORIGIN
sig_peptide
mat_peptide
Alignment Scores:
Pred. No.: 3 24e-63 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-589-870B-2 (1-183) x SASTRAVG (1-638)
QY 1 MetArgLysIleValAlaAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
DB 50 ATGGCGAAGATCGTCGTTGCACGATCGCGCTTCCCTGACACGCGTCCGATTACGGCC 109
QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaAlaGly 40
DB 110 AGCGCTTCGGAGAGCCCTCCCAAGACTCGAAGGCCCGAGTTCGGCCGCGAGCCGCGC 169
QY 41 IleThrGlyThrTyrTrpLysGlnIleuGlySerThrPheIleValThrAlaGlyAlaAsp 60
DB 170 ATCACCAGGACCTGTTACACACAGCTCGGCTGCACCTTCATCTAACCAGGAGCCGAC 229
QY 61 G1AlaLeuThrGlyThrTyrGluSerAlaValGlyAsnAlaGluSerArgTyrValLeu 80
DB 230 GCGCGCCGTGACCGGAACCTAGCAGAGTCGGCGCGCAAGCGGAGCGGCTAGCTCTG 289
QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
DB 290 ACCGGTGTGAGAGAGCGCCCGGCGCACCGAGCGAGCGGCGCCCTCGTTGAGCG 349
QY 101 ValAlaIlePlyAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGlyGlnTyr 120
DB 350 GTGGCTGGAGAAATACCTACCGCAAGCCCACTCCGGAGCACGATGAGGCGCCAGTAC 409

QY 121 ValGlyIAlaGluAlaArgIleAsnThrGlnTrpLeuThrSerGlyThrThrGly 140
DB 410 GTGGCGGCGCGGAGCGAGATCAACCCAGTGGCTGTGACTTCGGGACACACGAG 469
QY 141 AlaAsnAlaIlePlySerThrLeuValGlyHisAspThrPheThrLysValLysProSer 160
DB 470 GCCAAGCGCTGGAAGTCCACGCTGTGCGCCACGACACCTTCACCAAGTGAAGCCGTC 529
QY 161 AlaAlaSerIleAspAlaAlaLysAlaGlyValAlaAsnGlyAsnProLeuAspAla 180
DB 530 GCGGCTTCATCGACGCGCGCAAGAGCGCGCTCAACACCGCAACCCCTCGACGCC 589
QY 181 ValGlnGln 183
DB 590 GTTCAGCAG 598
RESULT 2
A93649
LOCUS A93649 638 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent EP0799890.
ACCESSION A93649
VERSION A93649.1 GI:6741838
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 638)
Mueller, R.D. and Deger, A.D.
Recombinant inactive core streptavidin mutants
Patent: EP 0799890-A 1 08-Oct-1997;
JOURNAL BOEHRINGER MANNHEIM GMBH (DE)
FEATURES
source location/Qualifiers
1..638
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
50..601
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB69587.1"
/db_xref="GI:6741838"
/translation="MRKIVAAIAVSLTIVSTASADPSKDSKQVSAARAGITGT
WYNQGSTFIVTAGDGLTGYESAVGNAESRYVLGRYDAPATDSGTALGMWIVA
KNNYRNHSAITWGGVGAERINTQWLTSSTTEANAKSTLVGHDTFTKVKPS
AASIDAKKAGVNNGNPLDAVQ"
50..121
sig_peptide
mat_peptide
BASE COUNT 115 a 244 c 193 g 86 t
ORIGIN
Alignment Scores:
Pred. No.: 3 24e-63 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-589-870B-2 (1-183) x A93649 (1-638)
QY 1 MetArgLysIleValAlaAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
DB 50 ATGGCGAAGATCGTCGTTGCACGATCGCGCTTCCCTGACACGCGTCCGATTACGGCC 109
QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaAlaGly 40
DB 110 AGCGCTTCGGAGAGCCCTCCCAAGACTCGAAGGCCCGAGTTCGGCCGCGAGCCGCGC 169
QY 41 IleThrGlyThrTyrTrpLysGlnIleuGlySerThrPheIleValThrAlaGlyAlaAsp 60
DB 170 ATCACCAGGACCTGTTACACACAGCTCGGCTGCACCTTCATCTAACCAGGAGCCGAC 229

QY 61 G1A1A1euthrG1yThrTg1uSerAlaValAG1yAenAlaG1uSerArG1yVal1leu 80
| | | | |
Db 230 GGGGCCCTGACCGGAACCTACAGAGTCGGCCGTGGCAAGCCGACAGCCGCTACGTCCTG 289
| | | | |
QY 81 ThrG1yArG1yThrAspSerAlaProAlaThrAspG1ySerG1yThra1A1eug1yTTPThr 100
| | | | |
Db 290 ACCGGTCGTACGACAGCCGCCGCCGACGAGCGGACCGGACCGGCTCGGTGGAGC 349
| | | | |
QY 101 ValAla1eThrLyAspAenThrArgAenAlaHisSerAlaThrThrTPSerg1yG1nTyr 120
| | | | |
Db 350 GTGGCTGGAAAGATACCTACCGCAAGCCGCTCCGACCACTGGAGCGGCGGACGATAC 409
| | | | |
QY 121 ValG1yG1yAlaG1uAlaArg1leAenThrG1nTTPLeuThrSerG1yThrThrG1u 140
| | | | |
Db 410 GTGGCGGGCCGACGAGCGAGATCAACACCACTGGCTGTACCTCCGGACACCGAG 469
| | | | |
QY 141 AlaAenAla1eThrLySerThrLeuValG1yHisAspThrPheThrLyVal1yProSer 160
| | | | |
Db 470 GCCAAGCCTGGAAGTCCAGCTGGTGGCCAGACCTTCACCAAGGTGAAGCCGTCC 529
| | | | |
QY 161 AlaAlaSer1leAspAla1Ala1yS1yAlaG1yValAenAenG1yAenProLeuAenAla 180
| | | | |
Db 530 GCCGCTTCATCGACGCGGGAAGGCGGCGTCAACAGCAACCGCTCGAGCGCC 589
| | | | |
QY 181 ValG1nG1n 183
| | | | |
Db 590 GTTCAGCAG 598
| | | | |
RESULT 3
LOCUS AR177252 638 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6312916.
ACCESSION AR177252
VERSION AR177252.1 GI:17919607
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 638)
AUTHORS Kopecky,E., Muller,R., Engh,R., Schmitt,U., Deger,A. and
Brandstetter,H.
TITLE Recombinant inactive core streptavidin mutants
JOURNAL Patent: US 6312916-A 1 06-NOV-2001;
FEATURES
source 1..638
location/Qualifiers
BASE COUNT 115 a 244 c 193 g 86 t
ORIGIN
/organism="unknown"
Alignment Scores:
Pred. No.: 3,24e-63 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-589-870B-2 (1-183) x AR177252 (1-638)
QY 1 MetArgLyS1leValAla1Ala1leAlaValSerLeuThrThrValSer1leThra1a 20
| | | | |
Db 50 ATGCGCAAGATCGTCTTGACGACATCGCGCTTCCCTGACCAAGGTCTCGATTACGGCC 109
| | | | |
QY 21 SerAlaSerAlaAspProSerLyAspSerLyAsp1G1nValSerAla1aG1uAlaG1y 40
| | | | |
Db 110 AGCGCTTCGGAGACCCCTCCAAAGACTCGAAGGCCAGGTCTCGGCCCGCGAGCGGCG 169
| | | | |
QY 41 1leThrG1yThrTTPThrAsnG1nLeuG1ySerThrPhe1leValThra1aG1yAlaAsp 60
| | | | |
Db 170 ATCACCGGACCTGTATCAACACGCTGGCTCGACTTCATCGTACCGCGGCGCGGAC 229
| | | | |
QY 61 G1yAla1euthrG1yThrTg1uSerAlaValAG1yAenAlaG1uSerArG1yVal1leu 80
| | | | |

Db 230 GGGGCCCTGACCGGAACCTACAGAGTCGGCCGTGGCAAGCCGACAGCCGCTACGTCCTG 289
| | | | |
QY 81 ThrG1yArG1yThrAspSerAlaProAlaThrAspG1ySerG1yThra1A1eug1yTTPThr 100
| | | | |
Db 290 ACCGGTCGTACGACAGCCGCCGCCGACGAGCGGACCGGACCGGCTCGGTGGAGC 349
| | | | |
QY 101 ValAla1eThrLyAspAenThrArgAenAlaHisSerAlaThrThrTPSerg1yG1nTyr 120
| | | | |
Db 350 GTGGCTGGAAAGATACCTACCGCAAGCCGCTCCGACCACTGGAGCGGCGGACGATAC 409
| | | | |
QY 121 ValG1yG1yAlaG1uAlaArg1leAenThrG1nTTPLeuThrSerG1yThrThrG1u 140
| | | | |
Db 410 GTGGCGGGCCGACGAGCGAGATCAACACCACTGGCTGTACCTCCGGACACCGAG 469
| | | | |
QY 141 AlaAenAla1eThrLySerThrLeuValG1yHisAspThrPheThrLyVal1yProSer 160
| | | | |
Db 470 GCCAAGCCTGGAAGTCCAGCTGGTGGCCAGACCTTCACCAAGGTGAAGCCGTCC 529
| | | | |
QY 161 AlaAlaSer1leAspAla1Ala1yS1yAlaG1yValAenAenG1yAenProLeuAenAla 180
| | | | |
Db 530 GCCGCTTCATCGACGCGGGAAGGCGGCGTCAACAGCAACCGCTCGAGCGCC 589
| | | | |
QY 181 ValG1nG1n 183
| | | | |
Db 590 GTTCAGCAG 598
| | | | |
RESULT 4
LOCUS AR204862 638 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6368813.
ACCESSION AR204862
VERSION AR204862.1 GI:21502296
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 638)
AUTHORS Reznik,G.O., Sano,T., Vajda,S., Smith,C. and Cantor,C.
TITLE Multifactor streptavidin
JOURNAL Patent: US 6368813-A 1 09-APR-2002;
FEATURES
source 1..638
location/Qualifiers
BASE COUNT 115 a 244 c 193 g 86 t
ORIGIN
/organism="unknown"
Alignment Scores:
Pred. No.: 3,24e-63 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-589-870B-2 (1-183) x AR204862 (1-638)
QY 1 MetArgLyS1leValAla1Ala1leAlaValSerLeuThrThrValSer1leThra1a 20
| | | | |
Db 50 ATGCGCAAGATCGTCTTGACGACATCGCGCTTCCCTGACCAAGGTCTCGATTACGGCC 109
| | | | |
QY 21 SerAlaSerAlaAspProSerLyAspSerLyAsp1G1nValSerAla1aG1uAlaG1y 40
| | | | |
Db 110 AGCGCTTCGGAGACCCCTCCAAAGACTCGAAGGCCAGGTCTCGGCCCGCGAGCGGCG 169
| | | | |
QY 41 1leThrG1yThrTTPThrAsnG1nLeuG1ySerThrPhe1leValThra1aG1yAlaAsp 60
| | | | |
Db 170 ATCACCGGACCTGTATCAACACGCTGGCTCGACTTCATCGTACCGCGGCGCGGAC 229
| | | | |
QY 61 G1yAla1euthrG1yThrTg1uSerAlaValAG1yAenAlaG1uSerArG1yVal1leu 80
| | | | |
Db 230 GGGGCCCTGACCGGAACCTACAGAGTCGGCCGTGGCAAGCCGACAGCCGCTACGTCCTG 289
| | | | |

QY 81 ThrGlyAgtYAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThr 100
 Db 290 ACCGGTCGTTAGACAGCCGCCCGGACCGAGCGGACCGCCCTCGGTTGACG 349
 QY 101 VALAATrPLyASnAntYrArGAsnAlaHisSerAlaThrThrTrpSerGlyGlnTrYr 120
 Db 350 GTGGCTTGAGAAATrACTrACCGCAAGCCCACTCGGCAACCGGAGCGGCGAGTAC 409
 QY 121 ValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrTrgIu 140
 Db 410 GTGGCGGGCGCGAGCGAGATCAACACCGCTGCTGCTGCTCGGACCGACCGAG 469
 QY 141 AlaAsnAlaTrpLysSerThrIleuValGlyHisAspThrPheThrIlyValLysProSer 160
 Db 470 GCCAAGCGCTGAGAGTCCAGCGCTGCGGCGACGACCTTCACCAAGGTGAAGCGCTCC 529
 QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
 Db 530 GCCGCTTCATCGACCGCGGAGAAAGCGCGGTCAACACGCGAACCGCTCGACCGCC 589
 QY 181 ValGlnGln 183
 Db 590 GTTCACGAG 598

RESULT 5
 AR217936 638 bp DNA linear PAT 25-SEP-2002
 LOCUS AR217936 Sequence 1 from patent US 6417331.
 DEFINITION AR217936
 ACCESSION AR217936
 VERSION AR217936.1 GI:23318240
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE
 1 (bases 1 to 638)
 Kopeckzki, E., Muller, R., Engh, R., Schmitt, U., Deger, A. and
 Brändsäter, H.
 TITLE Recombinant inactive core streptavidin mutants
 JOURNAL Patent: US 6417331-A 1 09-JUL-2002;
 FEATURES Location/Qualifiers
 source 1..638
 BASE COUNT 115 a 244 c 193 g 86 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,24e-63 Length: 638
 Score: 936.00 Matches: 183
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-589-870B-2 (1-183) x AR217936 (1-638)

QY 1 MetaGlySileValAlaAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
 Db 50 ATGGCCAGATCGTGGTTCAGCCATCGCGCTTCCCTGACCAACGCTCGATTACGGCC 109
 QY 21 SerAlaSerAlaAspProSerIlyAspSerIlyAlaGlnValSerAlaAlaGluAlaGly 40
 Db 110 AGCGCTTGCGGACGCCCTCCCAAGACTCGAAGGCGCGCTCGCGCGCGAGCGCGC 169
 QY 41 IleThrGlyThrTrpTrpYrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
 Db 170 ATCACCGGACCTGCTGACACCAAGCTCGCTGACCTTCATCTGACCGCGGCGCGAC 229
 QY 61 GlyAlaLeuThrGlyThrTrpGluSerAlaValaGlyAsnAlaGluSerArgTrpValLeu 80
 Db 230 GCGCGCTTGACCGGAGACTGAGAGTGGCGCTCGGCAACGCGGAGCGGCTAGCTCTG 289
 QY 81 ThrGlyAgtYAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThr 100

Db 290 ACCGGTCGTTAGACAGCCGCCCGGACCGAGCGGACCGCCCTCGGTTGACG 349
 QY 101 VALAATrPLyASnAntYrArGAsnAlaHisSerAlaThrThrTrpSerGlyGlnTrYr 120
 Db 350 GTGGCTTGAGAAATrACTrACCGCAAGCCCACTCGGCAACCGGAGCGGCGAGTAC 409
 QY 121 ValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrTrgIu 140
 Db 410 GTGGCGGGCGCGAGCGAGATCAACACCGCTGCTGCTGCTCGGACCGACCGAG 469
 QY 141 AlaAsnAlaTrpLysSerThrIleuValGlyHisAspThrPheThrIlyValLysProSer 160
 Db 470 GCCAAGCGCTGAGAGTCCAGCGCTGCGGCGACGACCTTCACCAAGGTGAAGCGCTCC 529
 QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
 Db 530 GCCGCTTCATCGACCGCGGAGAAAGCGCGGTCAACACGCGAACCGCTCGACCGCC 589
 QY 181 ValGlnGln 183
 Db 590 GTTCACGAG 598

RESULT 6
 AX057941 638 bp DNA linear PAT 17-JAN-2001
 LOCUS AX057941 Sequence 1 from Patent WO0075333.
 DEFINITION AX057941
 ACCESSION AX057941
 VERSION AX057941.1 GI:12310569
 KEYWORDS
 SOURCE Streptomyces avidinii
 ORGANISM Streptomyces avidinii
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Streptomycinae; Streptomycetaceae; Streptomyces.

REFERENCE
 1
 Goshorn, S.C., Graves, S.S., Schultz, J.E., Lin, Y., Sanderson, J.A. and
 Reno, J.M.
 TITLE Streptavidin expressed gene fusions and methods of use thereof
 JOURNAL Patent: WO 0075333-A 1 14-DEC-2000;
 FEATURES Location/Qualifiers
 source 1..638
 BASE COUNT 115 a 244 c 193 g 86 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,24e-63 Length: 638
 Score: 936.00 Matches: 183
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-589-870B-2 (1-183) x AX057941 (1-638)

QY 1 MetaGlySileValAlaAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
 Db 50 ATGGCCAGATCGTGGTTCAGCCATCGCGCTTCCCTGACCAACGCTCGATTACGGCC 109
 QY 21 SerAlaSerAlaAspProSerIlyAspSerIlyAlaGlnValSerAlaAlaGluAlaGly 40
 Db 110 AGCGCTTGCGGACGCCCTCCCAAGACTCGAAGGCGCGCTCGCGCGCGAGCGCGC 169
 QY 41 IleThrGlyThrTrpTrpYrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
 Db 170 ATCACCGGACCTGCTGACACCAAGCTCGCTGACCTTCATCTGACCGCGGCGCGAC 229
 QY 61 GlyAlaLeuThrGlyThrTrpGluSerAlaValaGlyAsnAlaGluSerArgTrpValLeu 80
 Db 230 GCGCGCTTGACCGGAGACTGAGAGTGGCGCTCGGCAACGCGGAGCGGCTAGCTCTG 289

Qy 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
|
|
|
Db 290 ACCGGTCGTTAGACAGCGCCCGGCGCACGAGCGGACCGCCCTCGTTGGAGC 349
|
|
|
Qy 101 ValAlaTyrPLeuAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGlyGlnTyr 120
|
|
|
Db 350 GTGGCTCGAAGAAATACCACTACCGCAACGCCCACTCCGCGACACGAGCGCCAGTAC 409
|
|
|
Qy 121 ValGlyGlyAlaGluAlaArgGlyLeuAsnThrGlnTyrPLeuLeuThrSerGlyThrThrGlu 140
|
|
|
Db 410 GTGGGGGGCGCGAGAGATCAACACCACTGCTCTGCTGCTCCGCGACACCGAG 469
|
|
|
Qy 141 AlaAsnAlaTyrPLeuSerThrLeuValGlyHisAspThrPheThrIleValIysProSer 160
|
|
|
Db 470 GCCAAGCGCTGGAAATCCACGCTGCTCGGCGCACGACCTTCCACCAAGTGAAGCCGTCC 529
|
|
|
Qy 161 AlaAlaSerIleLeuAlaAlaIleValIysValIleValIleAsnGlyIleAsnProLeuAla 180
|
|
|
Db 530 GCCGCTCCATCGACCGCGCGAAGAGCGCGGCTCAACACGCGAAGCCGCTCGACGCC 589
|
|
|
Qy 181 ValGlnGln 183
|
|
|
Db 590 GTTCAGCAG 598
|
|
|
RESULT 7
101349 638 bp ss-DNA linear PAT 21-MAY-1993
LOCUS 101349
DEFINITION Sequence 2 from Patent US 4839293.
ACCESSION 101349
VERSION 101349.1 GI:270135
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 638)
AUTHORS Cantor, C.R., Axel, R. and Argarana, C.
TITLE DNA encoding streptavidin, streptavidin produced therefrom, fused
polypeptides which include amino acid sequences present in
streptavidin and uses thereof
JOURNAL Patent: US 4839293-A 2 13-JUN-1989;
The Trustees of Columbia University in the City of New York; New
York, NY
FEATURES
source 1..638 location/Qualifiers
BASE COUNT 114 a 244 c 194 g 86 t
ORIGIN
Alignment Scores:
Pred. No.: 5.5e-63 Length: 638
Score: 933.00 Matches: 182
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 0
Query Match: 99.68% Indels: 0
Gaps: 0
US-09-589-870B-2 (1-183) x 101349 (1-638)
Qy 1 MetArgIysIleValAlaAlaAlaAlaAlaSerLeuThrThrValSerIleThrAla 20
|
|
|
Db 50 ATGGCGAAGATCGTGGTCGACGATCGCGCTTCCCTGACACCGGTCTCGATTACGGCC 109
|
|
|
Qy 21 SerAlaSerAlaAspProSerIleValAspSerIleValGlnValSerAlaAlaGluAlaGly 40
|
|
|
Db 110 ACGGCTTCGGAGAGACCCCTCCAGAGACTCGAGGCGCCGAGTCTCGGCCCGCCAGCGCCG 169
|
|
|
Qy 41 IleThrGlyThrTyrPLeuGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
|
|
|
Db 170 ATCCACCGGACCTCGTACCAACACGCTCGCTCATCTGACACCGCGCGCCGAC 229
|
|
|
Qy 61 GlyAlaLeuThrGlyThrTyrGlnSerAlaValGlyAsnAlaGluSerArgTyrValLeu 80
|
|
|
Db 230 GCGGCTCGACCGAAGCTACGAGTGGCGCTCGGACGCGAGAGCGCGTACGCTCGT 289
|
|
|

Qy 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
|
|
|
Db 290 ACCGGTCGTTAGACAGCGCCCGGCGCACGAGCGGACCGCCCTCGTTGGAGC 349
|
|
|
Qy 101 ValAlaTyrPLeuAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGlyGlnTyr 120
|
|
|
Db 350 GTGGCTCGAAGAAATACCACTACCGCAACGCCCACTCCGCGACACGAGCGCCAGTAC 409
|
|
|
Qy 121 ValGlyGlyAlaGluAlaArgGlyLeuAsnThrGlnTyrPLeuLeuThrSerGlyThrThrGlu 140
|
|
|
Db 410 GTGGGGGGCGCGAGAGATCAACACCACTGCTCTGCTGCTCCGCGACACCGAG 469
|
|
|
Qy 141 AlaAsnAlaTyrPLeuSerThrLeuValGlyHisAspThrPheThrIleValIysProSer 160
|
|
|
Db 470 GCCAAGCGCTGGAAATCCACGCTGCTCGGCGCACGACCTTCCACCAAGTGAAGCCGTCC 529
|
|
|
Qy 161 AlaAlaSerIleLeuAlaAlaIleValIysValIleValIleAsnGlyIleAsnProLeuAla 180
|
|
|
Db 530 GCCGCTCCATCGACCGCGCGAAGAGCGCGGCTCAACACGCGAAGCCGCTCGACGCC 589
|
|
|
Qy 181 ValGlnGln 183
|
|
|
Db 590 GTTCAGCAG 598
|
|
|
RESULT 8
57877 625 bp DNA linear BCT 30-OCT-1995
LOCUS 57877
DEFINITION streptavidin v1 [Streptomyces violaceus, Tu 2460, Genomic, 625 nt].
ACCESSION 57877
VERSION 57877.1 GI:1042193
KEYWORDS
SOURCE Streptomyces violaceus
ORGANISM Streptomyces violaceus
REFERENCE 1 (bases 1 to 625)
AUTHORS Bayer E.A., Kulik, T., Adar, R. and Witchek, M.
TITLE Close similarity among streptavidin-like, biotin-binding proteins
from Streptomyces
JOURNAL Biochim. Biophys. Acta 1263 (1), 60-66 (1995)
MEDLINE 95359204
PUBMED 7632734
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI gi169185] from the original journal article.
This sequence comes from Fig. 4.
FEATURES
source 1..625 location/Qualifiers
/organism="Streptomyces violaceus"
/mol_type="genomic DNA"
/db_xref="taxon:1936"
50..601
/gene="streptavidin v1, Sa v1"
50..601
/gene="streptavidin v1, Sa v1"
/note="This sequence comes from Fig. 5; Sa v1"
/codon_start=1
/product="streptavidin v1"
/protein_id="AAB35015.1"
/db_xref="GI:1042194"
/translation="MKRIKVAIAVSLTTSITASADPSKDAQVSAAEAGITGT
WKNYRNHSAHTTSGGVGTGRTAINTQMLTSTETANMKSTLVGHDFTKVPS
AASIDAKKAAVNNGNPDAVQO"
BASE COUNT 115 a 240 c 186 g 84 t
ORIGIN
Alignment Scores:
Pred. No.: 6.41e-63 Length: 625
Score: 932.00 Matches: 182
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.57% Indels: 0

DB: 1 Gaps: 0
US-09-589-870b-2 (1-183) x S78777 (1-625)
QY 1 MetArgLysIleValAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
Db 50 ATCCGCAAGATCGTGTGGAGCCATCGCGTTTCCCTGACCGGCTCGATTACCGCC 109
QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaIleAlaGly 40
Db 110 AGCGCTTCGGCAACCCCTCCAGACTCGAAGGCCAGGTCTCGGCCCGGAGCCGGC 169
QY 41 IleThrGlyThrTrpTyrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
Db 170 ATCCCGGCACTCGGTACCAACAGCTCGGCTCGACCTTCATCGTACCGCGGCGCCGAC 229
QY 61 GlnValLeuThrGlyThrTyrGlnSerAlaValAlaGlnAsnAlaGlnSerArgTyrValLeu 80
Db 230 GCGGCGCTGACCGGAACCTACGAGCTGCGCGTGGCAAGCGCCGAGAGCGGCTACGCTCG 289
QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThr 100
Db 290 ACCGCTGTACGACAGCGCGCCCGGACCGAGCGAGCGGACCGCCCTCGGCTGGACG 349
QY 101 ValAlaTrpLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyr 120
Db 350 GTGGCTGGAAGATTAACCAACCGCAACGCCCACTCCGACGACGAGCGGAGCGGCTAC 409
QY 121 ValGlyValAlaGlnAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThrGlu 140
Db 410 GTCCGGGACCGAGCGGAGATCAACACCGAGTGTGTGCTGCTCGGACCGACCGAG 469
QY 141 AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSer 160
Db 470 GCCAAGCGCTGGAGTCCAGCGTGTGGCCACGACCACTTCACCAAGGTGAAGCCGTCC 529
QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
Db 530 GCGCGCTTCATCGACCGCGGAGAAAGCGCGGTCAACACGCGAACCCGCTCGACGCC 589
QY 181 ValGlnGln 183
Db 590 GTTCAGCAG 598
RESULT 9
LOCUS S78782 625 bp DNA linear BCT 30-OCT-1995
DEFINITION streptavidin v2 [Streptomyces violaceus, Tu 2605, Genomic, 625 nt].
ACCESSION S78782
VERSION S78782.1 GI:1042195
KEYWORDS
SOURCE
ORGANISM Streptomyces violaceus
Streptomyces violaceus
Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
Streptomycetaceae; Streptomycetaceae; Streptomyces.
REFERENCE
AUTHORS 1 (bases 1 to 625)
TITLE Beyer, E.A., Kolik, T., Adar, R. and Wilchek, M.
JOURNAL Close similarity among streptavidin-like, biotin-binding proteins
MEDLINE from Streptomyces
PUBMED Biochim. Biophys. Acta 1263 (1), 60-66 (1995)
REMARK 7637734
Genbank scaff at the National Library of Medicine created this
entry [NCBI gisbseq 169186] from the original journal article.
This sequence comes from Fig. 4.
FEATURES
source
1..625
Location/Qualifiers
/organism="Streptomyces violaceus"
/mol_type="genomic DNA"
/db_xref="taxon:1936"
50..601
/gene="streptavidin v2, Sa v2"
50..601
CDS

/gene="streptavidin v2, Sa v2"
/note="This sequence comes from Fig. 5; Sa v2"
/codon_start=1
/product="streptavidin v2"
/protein_id="AAB35016.1"
/db_xref="GI:1042196"
translation="MRKIVAAIVSLTTVGITASASAPSKDSKAQAAVAEAGTGT
WYNOLGTFEIVTANADSLTGTSASVANSRYVLTGRVDSAPATDGGTALCMVTA
WKNVNRNHSATWISGVYVASSPARINTQMLITGTTAANMKSTLVGHDTFTKVPK
AASIDPAKAGVNNGNPLDAVQ"

BASE COUNT 117 a 244 c 184 g 80 t

ORIGIN

Alignment Scores:
Pred. No.: 3,04e-60 Length: 625
Score: 897.00 Matches: 174
Percent Similarity: 96.72% Conservative: 3
Best Local Similarity: 95.08% Mismatches: 6
Query Match: 95.83% Indels: 0
DB: 1 Gaps: 0

US-09-589-870b-2 (1-183) x S78782 (1-625)
QY 1 MetArgLysIleValAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
Db 50 ATCCGCAAGATCGTGTGGAGCCATCGCGTTTCCCTGACCGGCTCGATTACCGCC 109
QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaIleAlaGly 40
Db 110 AGCGCTTCGGCAACCCCTCCAGACTCGAAGGCCAGGTCTCGGCCCGGAGCCGGC 169
QY 41 IleThrGlyThrTrpTyrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
Db 170 ATCCCGGCACTCGGTACCAACAGCTCGGCTCGACCTTCATCGTACCGCGGCGCCGAC 229
QY 61 GlnValLeuThrGlyThrTyrGlnSerAlaValAlaGlnAsnAlaGlnSerArgTyrValLeu 80
Db 230 GCGGCGCTGACCGGAACCTACGAGCTGCGCGTGGCAAGCGCCGAGAGCGGCTACGCTCG 289
QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThr 100
Db 290 ACCGCTGTACGACAGCGCGCCCGGACCGAGCGGACCGCCCTCGGCTGGACG 349
QY 101 ValAlaTrpLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyr 120
Db 350 GTGGCGTGAAGAAACCAACCGCAACGCCCACTCCGACGACGAGCGGAGCGGCTAC 409
QY 121 ValGlyValAlaGlnAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThrGlu 140
Db 410 GTCCCGGCTTCGAGCGCGGATCAACACCGAGTGTGTGCTGCTCGGCGGACCGCGC 469
QY 141 AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSer 160
Db 470 GCCAAGCGCTGGAGTCCAGCGTGTGGCCACGACCACTTCACCAAGGTGAAGCCGTCC 529
QY 161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
Db 530 GCGCGCTTCATCGACCGCGGAGAAAGCGCGGTCAACACGCGAACCCGCTCGACGCC 589
QY 181 ValGlnGln 183
Db 590 GTCCAGCAG 598
RESULT 10
LOCUS AX057945 1239 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 5 from Patent WO0075333.
ACCESSION AX057945
VERSION AX057945.1 GI:12310571
KEYWORDS
SOURCE Streptomyces avidinii
ORGANISM Streptomyces avidinii
Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;

REFERENCE	Streptomyces; Streptomycetaceae; Streptomyces.			
AUTHORS	1 Goehorn, S. C., Graves, S. S., Schultz, J. E., Lin, Y., Sanderson, J. A. and Reno, J. M.			
TITLE	Streptavidin expressed gene fusions and methods of use thereof			
JOURNAL	Patent: WO 0075333-A 5 14-DEC-2000; NEORX CORPORATION (US)			
FEATURES	Location/Qualifiers			
source	1..1239 /organism="Streptomyces avidinii" /mol_type="genomic DNA" /db_xref="taxon:1895"			
BASE COUNT	270 a	392 c	356 g	221 t
ORIGIN				
Alignment Scores:				
Pred. No.:	9,67e-57	Length:	1239	
Score:	855.50	Matches:	165	
Percent Similarity:	98.83%	Conservative:	4	
Best Local Similarity:	96.48%	Mismatches:	1	
Query Match:	91.40%	Indels:	1	
Di:	6	Gaps:	1	
US-09-589-870B-2 (1-183) x AX057945 (1-1239)				
OY	14	ThrThrValSerIleThrAla----	SerAlaSerAlaAspProSerIleAspSerIleAla	32
Db	724	ACCAGGCTACCGTAGCTCGGCTCTGTTGGCAGACCCCTCAAGATCGAAGCC		783
OY	33	GlnValSerAlaAlaGluAlaGlyYlleThrGlyThrTAPYrAsnGlnIleGlySerThr		52
Db	784	CAGGCTCGCGCCGAGCGCGGATCCAGGACCTGGTACACAGCTGGCTTCACC		843
OY	53	PheIleValIThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTYrGlySerAlaValGly		72
Db	844	TTCATCGTAGACCGCGCGCGCGGCGCCCTGACCGGAACCTAGAGTCGCCGTGGC		903
OY	73	AsnAlaGlySerArgTYrValIleuThrGlyArgTYrAspSerAlaProAlaThrAspGly		92
Db	904	AACGCGAGACCCGCTACGCTCGTACCGCGGTACGACAGCGCGCGCACCGAGCGC		963
OY	93	SerGlyThrAlaIleGlyTYrThrValAlaIleTrpIleAsnArgTYrArgAsnAlaIleSer		112
Db	964	AGCGGACCGCCCTCGTGGACGGTGGCCCTGGAAGATTAATCCGCAAGCCCACTCC		1023
OY	113	AlaThrThrTYrTPSerGlyGlnTYrValGlyGlyAlaGluAlaArgIleAsnThrGlnTrp		132
Db	1024	CGGACCACTGGAGCGGCACATGCTCGCGCGCCCGGAGGAGATCAACACCACTGG		1083
OY	133	LeuIleuThrSerGlyTYrThrGluAlaAsnAlaIleTrpIleSerThrIleuValGlyIleAsp		152
Db	1084	CTGCTGACCTCGCGACCAACGAGGCCAACCCCTGGAAGTCCACGCTGTGGCCAGAC		1143
OY	153	ThrPheThrIleValIleProSerAlaAlaSerIleAspAlaAlaIleValGlyAlaGlyVal		172
Db	1144	ACCTTCACCAAGGTGAACCGCTCGCGCCCTTCATCGACCGCGGGAAGGCGCGGCTC		1203
OY	173	AsnAsnGlyAsnProIleuAspAlaValGlnGln		183
Db	1204	AACAAAGGCAACCCGCTGACCGCCGTTACGACG		1236
RESULT 11				
AX057943	1612 bp	DNA	linear	PAT 17-JAN-2001
LOCUS				
DEFINITION	Sequence 3 from Patent WO0075333.			
ACCESSION	AX057943			
VERSION	AX057943.1 GI:12310570			
KEYWORDS				
SOURCE	Streptomyces avidinii			
ORGANISM	Streptomyces avidinii			
REFERENCE	Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.			

FEATURES	source	location/Qualifiers
BASE COUNT	353 a	493 c 459 g 307 t
ORIGIN	/organism="Streptomyces avidinii" /mol_type="genomic DNA" /db_xref="taxon:1895"	
Alignment Scores:	1.54e-56	Length: 1612
Pred. No.:	854.50	Matches: 166
Score:	97.67%	Conservative: 2
Percent Similarity:	96.51%	Mismatches: 3
Best Local Similarity:	91.29%	Indels: 1
Query Match:	6	Gaps: 1
DB:	US-09-589-870B-2 (1-183) x AX057943 (1-1612)	
Qy	12 SerleuThrValSerIleThrAlaSerAlaSerAlaAspProSerIleAspSerIle	31
Db	1091 ACCTTACTGACCGGAGAC--TCGGCTCTGGTTGGCAACCCCTCCAGACTCCAG	1147
Qy	32 AlaGlnValSerAlaAlaGlnAlaGlyIleThrGlyThrTrpIleAsnGlnLeuGlySer	51
Db	1148 GCCACGCTCTCGCGCCCGAGCGCCGATACCGGACCTGTGTAACAACGCTCGCTCG	1207
Qy	52 ThrPheIleValThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTyrgIleSerAlaVal	71
Db	1208 ACCTTCATCGAGACCGCGGGCGGAGCGGCCCTGACCGGAACCTACGATCGCGCCTC	1267
Qy	72 GlyAsnAlaGlnSerAlaGlyIleValLeuThrGlyAlaArgTyIleAspSerAlaProAlaThrAsp	91
Db	1268 GGCACACCCGAGACCCCTCAGCTCTGACCGGTCTTACGACAGCGCCCGCCACCGAC	1327
Qy	92 GlySerGlyThrAlaLeuGlyIleTrpThrValAlaTrpIleAsnAsnTyIleArgAlaAlaHis	111
Db	1328 GGCAGCGGCACCGCCCTCGTTGGACGCGTGGCTCGGAAGAAATACATCCGACACGCCAC	1387
Qy	112 SerAlaThrThrTrpSerGlyGlnTyIleValGlyIleAlaGlnAlaArgIleAsnThrGln	131
Db	1388 TCCGCGACCAAGTGGACGGCCAGTACGTGCGGGCGCGGCGGAGATCAACACCCAG	1447
Qy	132 TrpLeuLeuThrSerGlyIleThrGlnAlaAsnAlaTrpIleSerThrIleValGlyHis	151
Db	1448 TGGCTGCTGACCTCCGCGACACACCGAGGCCAACGCTCGAAGTCCAGCTGTGCGGCAC	1507
Qy	152 AspThrPheThrIleValIleValProSerAlaAlaSerIleAspAlaAlaIleValIleValGly	171
Db	1508 GACACCTTACCAAGGAGGAGCGGTCCGCGCTTCATCCAGCGGCGAAGAGGCGCGC	1567
Qy	172 ValAsnAsnGlyAsnProLeuAspAlaValAlaGlnGln	183
Db	1568 GTCAACAACGGCAACCCGCTCGACGCCGTTCAACAG	1603
RESULT 12		
AX057947	1280 bp	DNA linear PAT 17-JAN-2001
LOCUS	Sequence 7	from Patent WO0075331.
DEFINITION	AX057947	
ACCESSION	AX057947.1	GI:12310572
VERSION		
KEYWORDS		
SOURCE	Streptomyces avidinii	
ORGANISM	Streptomyces avidinii	
REFERENCE	Bacteria, Actinobacteria; Actinobacteridae; Actinomycetales;	
AUTHORS	Streptomycineae; Streptomyceaceae; Streptomyces.	
	1 Goshorn, S.C., Graves, S.S., Schultz, J.E., Lin, Y., Sanderson, J.A. and	
	Reno, J.M.	

TITLE Streptavidin expressed gene fusions and methods of use thereof
JOURNAL Patent: WO 0075333-A 7 14-DEC-2000;
NEORX CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1280
/organism="Streptomyces avidinii"
/mol_type="genomic DNA"
/db_xref="taxon:1895"

BASE COUNT 267 a 397 c 388 g 228 t
ORIGIN

Alignment Scores:
Pred. No.: 4,48e-56 Length: 1280
Score: 847.00 Matches: 162
Percent Similarity: 98.79% Conservative: 1
Best Local Similarity: 98.18 Mismatches: 2
Query Match: 90.49% Indels: 0
DB: 6 Gaps: 0

US-09-589-870b-2 (1-183) x AX057947 (1-1280)

QY 19 ThrAlaSerAlaSerAlaSPProSerLyAspSerLyAlaGlnValSerAlaGln 38
Db 777 TCTGGCTCTGCTTGGCGACAGCCCTCCAGAGACTCGAAGGCCAGGCTCTGGCCCGCGAG 836
QY 39 AlaGlyIleThrGlyThrTPTyAsnGlnLeuGlySerThPheIleValThrAlaGly 58
Db 837 GCGGCGATCACCGGCACCTGTAACAACACAGCTCGGCTGACCTTCATCTGACCGCGGCG 896
QY 59 AlaAspGlyAlaLeuThrGlyThrTPTyGlySerAlaValAlaGlyAsnAlaGlnSerArgTy 78
Db 897 GCGGAGGCGCCCTGACCGGAACCTACGAGTCGGCCGTGGCAACCGCGAGCGCTAC 956
QY 79 ValLeuThrGlyArgTyAspSerAlaProAlaThrAspGlySerGlyThraIleuGly 98
Db 957 GTCCGACCGGCTGTTACGACAGCGCCCGGCCACCGAGCGAGCGGCGCCCTCGGT 1016
QY 99 TPThrValAlaIleTPlyAsnAsnTPTyArgAsnAlaHisSerAlaThrThrTPSerGly 118
Db 1017 TGGAGCGTGGCTCGGAAGATTAACCTACCGCAAGCCCACTCGCGCACGCTGAGCGGCG 1076
QY 119 GlnTPyValGlyValAlaGlnAlaArgIleAsnThrGlnTPleuLeuThrseryGlyThr 138
Db 1077 CAGTACGTGGCGGCCCGAGCGAGATCAACACCACTGCTGCTGACCTCCGCGCAC 1136
QY 139 ThrGlnAlaAsnAlaTPlySerThrLeuValGlyHisAspThrPheThrLyValLys 158
Db 1137 ACCGAGCCCAAGCCCTGGAACTCCAGCTGCTGCGCCACGACACCTTCACCAAGGTGAAG 1196
QY 159 ProSerAlaAlaSerIleAspAlaAlaLysLysAlaGlyValAlaAsnGlnYAsnProLeu 178
Db 1197 CCGTCGCGCGCTCCATCGACGCGGAGAAAGCGGCGCTCAACACGCAACCGGCTC 1256
QY 179 AspAlaValGlnGln 183
Db 1257 GAGCGCGTTCAGCAG 1271

RESULT 13
A39565 1356 bp DNA linear PAT 05-MAR-1997
LOCUS A39565
DEFINITION Sequence 4 from Patent WO9415644.
ACCESSION A39565
VERSION A39565.1 GI:2295847
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 1356)
AUTHORS Epenetos,A.A., Spooner,R.A. and Deonaraiah,M.
TITLE COMPOUNDS FOR TARGETING
JOURNAL Patent: WO 9415644-A 4 21-JUL-1994;
IMP CANCER RES TECH (GB)
COMMENT Other publication GB 2289679 951129.

FEATURES Location/Qualifiers
source 1..1356
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
40..1347
/note="unnamed protein product"
/codon_start=1
/protein_id="CA02449.1"
/db_xref="GI:2295848"
/translation="MKYLLPTAAGLLILAAOPAAVOYLQOPGAELVPGASVYLSCKASGTFPSYMMHWKORPARGLEIGIDPNSGCTKYNKFKSKATLTVKPSSTAYMQLSLSSEDAVYICARVDYGSYFVWQGTIVTSVSGSGSGSGSGSGSGQAVVTOESALTPSPGETVTLICRSSTGAVTTSNANWQEPDHLFTGLIGTNNRAPGVPARFSGSLTIGDKAALITIGAOTEDEAIFYCALMYSNHWFGGGLTKLTVGLEAPAAPAPDSKSKQAVAAEAGITGWNOGSEFFIVTAGADGALVTGYESAAGNARSRYVLTGRYDAPATDGGSTALGWTVMKNNRNHNSHTWSGVVGAAEARIINQWLLTSGTTEANMKSTLVGHDITFKVPSAASIDAAKAGVNNGNPLDANVQ"

BASE COUNT 332 a 357 c 361 g 306 t
ORIGIN

Alignment Scores:
Pred. No.: 5,69e-56 Length: 1356
Score: 846.00 Matches: 162
Percent Similarity: 99.39% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 1
Query Match: 90.38% Indels: 0
DB: 6 Gaps: 0

US-09-589-870b-2 (1-183) x A39565 (1-1356)

QY 20 AlaSerAlaSerAlaAspProSerLyAspSerLyAlaGlnValSerAlaGlnAla 39
Db 853 GCTGCGGACCTCGAGACCCCTCCAGAGACTCAAAAGCTCAGGTTCTGACGCCAAGCT 912
QY 40 GlyIleThrGlyThrTPTyAsnGlnLeuGlySerThPheIleValThrAlaGlyVala 59
Db 913 GGTATCACTGGACACTGGATTAACCACTGGGGCTGACTTCATTGTGACCGCTGGTGG 972
QY 60 AspGlyAlaLeuThrGlyThrTPTyGlySerAlaValAlaGlyAsnAlaGlnSerArgTyVal 79
Db 973 GACGAGCTCTGACCTCGGACCTACGATCTCGGTGTAACGCAAGTCCCGCTACGTA 1032
QY 80 LeuThrGlyArgTyAspSerAlaProAlaThrAspGlySerGlyThraIleuGlyTyr 99
Db 1033 CTGACTGGCCGCTTATGACTTCGACCTGCCACCGATGCTCTGGTAACCGCTCGGCTGG 1092
QY 100 ThrValAlaIleTPlyAsnAsnTPTyArgAsnAlaHisSerAlaThrThrTPSerGlyGln 119
Db 1093 ACTGGGCTTGGAAACCACTATCTGTAATCGGACAGCGCACCTACGTGGCTGGCCAA 1152
QY 120 TyrValGlyValAlaGlnAlaArgIleAsnThrGlnTPleuLeuThrseryGlyThrThr 139
Db 1153 TACGTTGGCGGCTCGAAGCTCTGATCAACCTACAGTGGCTGTTAACATCCGGGCACTACC 1212
QY 140 GlnAlaAsnAlaIleTPlySerThrLeuValGlyHisAspThrPheThrLyValLysPro 159
Db 1213 GAAGCAATGATGAGAAATTCACACTAGTAGTCACTGACACCTTTACCAAGATTAAAGCT 1272
QY 160 SerAlaAlaSerIleAspAlaAlaLysLysAlaGlyValAlaAsnGlnYAsnProLeuAsp 179
Db 1273 TCTGCTGCTACATTGATGTGCCCAAGAACGAGCGTAAACACAGGTAAACCTCTTAGAC 1332
QY 180 AlaValGlnGln 183
Db 1333 GCTGTTACGAA 1344

RESULT 14
A93152 1356 bp DNA linear PAT 22-JAN-2000
LOCUS A93152
DEFINITION Sequence 4 from Patent EP0815872.
ACCESSION A93152
VERSION A93152.1 GI:6741540

KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1356)
AUTHORS Epenetos, A.A. and Deonaraín, M.
TITLE Compounds for targeting
JOURNAL Patent: EP 0815872-A 4 07-JAN-1998;
IMP CANCER RES TECH (GB)
FEATURES
source Location/Qualifiers
1..1356
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
40..1347
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB69540.1"
/db_xref="GI:6741541"
/translation="MKYLLPTAAAGLLLLAOPAMNQVLOQPGAEIVKSGAVKLSK
KAGYFTFSYMMHMKVORRGLEWIGRIDPSGCKVKEKSKATLTVDPSSSTAY
MGLSLETSRDSAVYICARYDYGSIFDYWGCTTVTVSSGGSGSGSGSGSGSOAV
VTQESALITTSPEETVLTICRSTGAVTTSNVANWQEKPDHFTGLIGSTNNRPGVP
ARFSGSLIDKALITTTGAQTDEALFYCALWYSHMWFGGCTKLTVLGLAPAPAPA
DSKDSKAQVSAEAGITGTWYQLGSTIVTAGAGALTGYESAAGVABSRVYLTV
RYDSAPATDGGSGTALGWTVAAMKNYNNAHSATTWGSGQYVGALEARINTQWLLTSGTTE
AAMKSTLVGHDTFTKVRPSAASIDAKKAGVNNNGRPDAVQ"

BASE COUNT 332 a 357 c 361 g 306 t
ORIGIN

Alignment Scores:
Pred. No.: 5,698-56 Length: 1356
Score: 846.00 Matches: 162
Percent Similarity: 99.39% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 1
Query Match: 90.38% Indels: 0
DB: Gaps: 0

US-09-589-870b-2 (1-183) x A93152 (1-1356)

QY 20 AAlaSerAlaSerAlaApProSerLyAspSerLyAlaGlnValSerAlaAlaGluAla 39
DB 853 GCTGCCGACCTGCACACCCGTCACAGACTCCAAAGCTCAGGTTCTGCAGCCGAACT 912

QY 40 GlyIleThrGlyThrTrpTyraGlnLeuGlySerThrPheIleValThrAlaGlyAla 59
DB 913 GGATCTACCTGGCACCCTGGTATACCAACTGGGGTCCAGCTTCACTGTGACCCCTGGTGG 972

QY 60 AapGlyAlaLeuThrGlyThrTyraGlnLeuGlySerThrPheIleValThrAlaGlyAla 79
DB 973 GACGGAGCTCTGACTGACCTGACCTTACGAACTGCGGTTGGTAACGCAAACTCCGCTACGTA 1032

QY 80 LeuThrGlyArgTyraPserAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyr 99
DB 1033 CTGACTGGCGGCTTATGACTCTGCACCTGCACGAGTGGCTGTGATCCCTCGGGCTGG 1092

QY 100 ThrValAlaATrPlyAsnAsnThrTyraGlnAlaHisSerAlaThrThrTrpSerGlyGln 119
DB 1093 ACTGGGCTTGGAAAAACAATCTGTAATGGCGCACAGGCCCACTACGTGGCTGGCCAA 1152

QY 120 TyrValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThr 139
DB 1153 TACGTTGGCGGCTGAGGCTCGTATCAACCTCAGTGGCTGTAAACATCCGGCACTAAC 1212

QY 140 GluAlaAsnAlaATrPlySerThrLeuValGlyHisAspThrPheThrLyValIysPro 159
DB 1213 GAAGCGAATGATGAAATCGACACTAGTACGATGACACCTTTACCAAGTTAAAGCT 1272

QY 160 SerAlaAlaSerIleAspAlaAlaIleValGlyValAsnAsnGlyAsnProLeuAsp 179
DB 1273 TCTGCTGTACGATTTGATGCTGCCAAGAAAGCGGCTAAACACGTAACCTCTAGAC 1332

QY 180 AlaValGlnGln 183

DB 1333 GCTGTTACGCAA 1344

RESULT 15
LOCUS AR082490 1356 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 4 from patent US 5973116.
ACCESSION AR082490
VERSION AR082490.1 GI:10009216
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1356)
AUTHORS Epenetos, A. Antoniou, J. Spooner, R. Anthony, and Deonaraín, M.
TITLE Compounds for targeting
JOURNAL Patent: US 5973116-A 4 26-OCT-1999;
FEATURES
source Location/Qualifiers
1..1356
/organism="unknown"
BASE COUNT 332 a 357 c 361 g 306 t
ORIGIN

Alignment Scores:
Pred. No.: 5,698-56 Length: 1356
Score: 846.00 Matches: 162
Percent Similarity: 99.39% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 1
Query Match: 90.38% Indels: 0
DB: Gaps: 0

US-09-589-870b-2 (1-183) x AR082490 (1-1356)

QY 20 AAlaSerAlaSerAlaApProSerLyAspSerLyAlaGlnValSerAlaAlaGluAla 39
DB 853 GCTGCCGACCTGCACACCCGTCACAGACTCCAAAGCTCAGGTTCTGCAGCCGAACT 912

QY 40 GlyIleThrGlyThrTrpTyraGlnLeuGlySerThrPheIleValThrAlaGlyAla 59
DB 913 GGATCTACCTGGCACCCTGGTATACCAACTGGGGTCCAGCTTCACTGTGACCCCTGGTGG 972

QY 60 AapGlyAlaLeuThrGlyThrTyraGlnLeuGlySerThrPheIleValThrAlaGlyAla 79
DB 973 GACGGAGCTCTGACTGACCTGACCTTACGAACTGCGGTTGGTAACGCAAACTCCGCTACGTA 1032

QY 80 LeuThrGlyArgTyraPserAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyr 99
DB 1033 CTGACTGGCGGCTTATGACTCTGCACCTGCACGAGTGGCTGTGATCCCTCGGGCTGG 1092

QY 100 ThrValAlaATrPlyAsnAsnThrTyraGlnAlaHisSerAlaThrThrTrpSerGlyGln 119
DB 1093 ACTGGGCTTGGAAAAACAATCTGTAATGGCGCACAGGCCCACTACGTGGCTGGCCAA 1152

QY 120 TyrValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThr 139
DB 1153 TACGTTGGCGGCTGAGGCTCGTATCAACCTCAGTGGCTGTAAACATCCGGCACTAAC 1212

QY 140 GluAlaAsnAlaATrPlySerThrLeuValGlyHisAspThrPheThrLyValIysPro 159
DB 1213 GAAGCGAATGATGAAATCGACACTAGTACGATGACACCTTTACCAAGTTAAAGCT 1272

QY 160 SerAlaAlaSerIleAspAlaAlaIleValGlyValAsnAsnGlyAsnProLeuAsp 179
DB 1273 TCTGCTGTACGATTTGATGCTGCCAAGAAAGCGGCTAAACACGTAACCTCTAGAC 1332

QY 180 AlaValGlnGln 183
DB 1333 GCTGTTACGCAA 1344

Search completed: October 28, 2003, 01:30:45
Job time : 3298 sec

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 28, 2003, 01:30:48 ; Search time 248 Seconds
(without alignments)
1978.910 Million cell updates/sec

Title: US-09-589-870B-2

Perfect score: 936
Sequence: 1 MKRIYVAIAVSLTYSITA.....IDAKKAGVNGNPLDAVQ 183

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlh
-Q/cg2_1/USPTO.epool/US09589870/runat_27102003_104450_16725/app.query.fasta_1.327
-DB=Published.Applications.NA -QFMT=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blossum62
-TRANS-human40.cgi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NOR=ext -HEARSIZE=500 -MILEN=0
-MAXLEN=200000000 -USER=US09589870.QCGN_1_1221@runat_27102003_104450_16725
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA.*

1:	/cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3:	/cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4:	/cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5:	/cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6:	/cg2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8:	/cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9:	/cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10:	/cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11:	/cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
12:	/cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13:	/cg2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14:	/cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
15:	/cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16:	/cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17:	/cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	100.0	638	12	US-10-244-821-1 Sequence 1, Appli

2	936	100.0	638	14	US-10-013-173-1	Sequence 1, Appli
3	936	100.0	638	14	US-10-150-762-1	Sequence 1, Appli
4	855.5	91.4	1239	12	US-10-244-821-5	Sequence 5, Appli
5	855.5	91.4	1239	14	US-10-013-173-5	Sequence 5, Appli
6	855.5	91.4	1239	14	US-10-150-762-5	Sequence 5, Appli
7	854.5	91.3	1614	12	US-10-244-821-3	Sequence 3, Appli
8	854.5	91.3	1614	14	US-10-013-173-3	Sequence 3, Appli
9	854.5	91.3	1614	14	US-10-150-762-3	Sequence 3, Appli
10	847	90.5	1280	12	US/10/244	Sequence 7, Appli
11	847	90.5	1280	14	US/10/013	Sequence 7, Appli
12	847	90.5	1280	14	US/10/150	Sequence 7, Appli
13	847	90.5	1271	12	US-10-244-821-87	Sequence 87, Appli
14	847	90.5	1467	12	US-10-244-821-48	Sequence 48, Appli
15	847	90.5	1467	14	US-10-013-173-48	Sequence 48, Appli
16	847	90.5	1467	14	US-10-150-762-48	Sequence 48, Appli
17	834	89.1	498	10	US-09-117-447-7	Sequence 7, Appli
18	834	89.1	1266	10	US-09-938-2708-2	Sequence 2, Appli
19	687.5	73.5	1173	12	US-10-075-947A-4	Sequence 4, Appli
20	682	72.9	1176	12	US-10-075-947A-3	Sequence 3, Appli
21	110.5	11.8	3381	11	US-09-952-267-6	Sequence 6, Appli
22	105.5	11.3	3018	12	US-10-193-764-40	Sequence 40, Appli
23	105.5	11.3	3036	12	US-10-193-764-38	Sequence 38, Appli
24	102.5	11.0	3331	9	US-09-864-761-19481	Sequence 19481, A
25	101.5	10.8	3300	9	US-09-379-931-6	Sequence 6, Appli
26	101.5	10.8	3300	12	US-10-223-597-6	Sequence 6, Appli
27	99.5	10.6	3108	12	US-10-193-764-58	Sequence 68, Appli
28	99.5	10.6	4837	12	US-10-193-764-56	Sequence 66, Appli
29	99.5	10.6	4937	13	US-10-092-880-3	Sequence 3, Appli
30	99.5	10.6	7407	12	US-10-246-830-3	Sequence 3, Appli
31	99.5	10.6	9323	13	US-10-092-880-6	Sequence 6, Appli
32	98	10.5	3349	11	US-09-952-267-2	Sequence 2, Appli
33	97.5	10.4	489	9	US-09-864-761-27528	Sequence 27528, A
34	96.5	10.3	1089	14	US-10-156-761-3313	Sequence 3313, Ap
35	96.5	10.3	2247	11	US-09-272-975-3	Sequence 3, Appli
36	96.5	10.3	7507	11	US-09-272-975-1	Sequence 1, Appli
37	96.5	10.3	902608	14	US-10-156-761-1	Sequence 1, Appli
38	95	10.1	902608	14	US-10-156-761-1	Sequence 1, Appli
39	94	10.0	801	10	US-09-974-300-6560	Sequence 6560, Ap
40	93.5	10.0	896	12	US-10-101-510-732	Sequence 732, Ap
41	93.5	10.0	1329	14	US-10-156-761-6352	Sequence 6352, Ap
42	92.5	9.9	825	14	US-10-156-761-7318	Sequence 7318, Ap
43	92.5	9.9	3711	11	US-09-769-787-321	Sequence 321, App
44	92.5	9.9	25165	14	US-10-114-170-39	Sequence 39, Appli
45	92	9.8	461	9	US-09-864-761-10890	Sequence 10890, A

ALIGNMENTS

RESULT 1
US-10-244-821-1
Sequence 1, Application US/10244821
Publication No. US20030143233A1
GENERAL INFORMATION:
APPLICANT: Goshorn, Stephen Charles
APPLICANT: Graves, Scott Stoll
APPLICANT: Schultz, Joanne Elaine
APPLICANT: Lin, Yukang
APPLICANT: Sandersen, James Allen
APPLICANT: Reno, John M.
APPLICANT: Dearsteyne, Erica A.
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
METHODS OF USE THEREOF
FILE REFERENCE: 690022.547C3
CURRENT APPLICATION NUMBER: US/10/244.821
CURRENT FILING DATE: 2002-09-16
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 638
TYPE: DNA
ORGANISM: Streptomyces avidinii
US-10-244-821-1

Alignment Scores:

Pred. No.: 1,92e-101 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-589-870B-2 (1-183) x US-10-244-821-1 (1-638)

QY 1 MetArgLySLeIValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
DB 50 ATGGGAGAGTCGTCGTCGAGCCATCGCGTTCCCTGACCAACGCGTCTCGATTACGGCC 109
QY 21 SerAlaSerAlaAspProSerIleAspSerIleValSerAlaGlnValSerAlaGlnValaGly 40
DB 110 AGGCGCTTGGCAGACCCCTCCAGAGACTCGAAGGCCAGGTCTCGGCCCGCCAGGCCGCG 169
QY 41 IleThrGlyThrTyrTyrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
DB 170 ATCACCGGACCTGCTACCAACAGCTCGGCTCGACCTTCATCGTACCGCGGCGCGAC 229
QY 61 GlyAlaLeuThrGlyThrTyrGlnSerAlaValGlyAsnAlaGlnSerArgTyrValLeu 80
DB 230 GGGCGCCCTAACCGAACCCTACGAGTCGCGCGGACGCGGACGCGCGCTACGCTCG 289
QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
DB 290 ACCGCTGTTACGACAGCGCCCGCGCCACGAGCGGACGCGCGCTCGGTTAGAG 349
QY 101 ValAlaTyrPlyAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGlyGlnTyr 120
DB 350 GTGGCTCGAAGATTAACCTACCGCAACGCCCACTCCGACCAACGCGGAGCGCGCTAC 409
QY 121 ValGlyGlyAlaGlnAlaArgIleAsnThrGlnTyrLeuLeuThrSerGlyThrThrGln 140
DB 410 GTGGCGGCGCGGAGCGGAGATCAACACCGAGTGTCTCTGCTCCGCGACCAACCGAG 469
QY 141 AlaAsnAlaTyrPlySerThrLeuValGlyHisAspThrPheThrIleValIleProSer 160
DB 470 GCCAAGCCTGGAAGTCCAGCGCTGCGCGCCACGACCTTCACCAAGTGAAGCGCTCC 529
QY 161 AlaAlaSerIleAspAlaAlaIleValIleValIleValIleValIleValIleValIle 180
DB 530 GCCGCTTCATCGACGCGCGGAGAAAGCGCGCTCAACACGGAACCGCTCGACGCC 589
QY 181 ValGlnGln 183
DB 590 GTTCAGCAG 598

RESULT 2

US-10-013-173-1
; Sequence 1, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013.173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Streptomyces avidinii
US-10-013-173-1

Alignment Scores:

Pred. No.: 1,92e-101 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-589-870B-2 (1-183) x US-10-013-173-1 (1-638)

QY 1 MetArgLySLeIValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
DB 50 ATGGGAGAGTCGTCGTCGAGCCATCGCGTTCCCTGACCAACGCGTCTCGATTACGGCC 109
QY 21 SerAlaSerAlaAspProSerIleAspSerIleValSerAlaGlnValSerAlaGlnValaGly 40
DB 110 AGGCGCTTGGCAGACCCCTCCAGAGACTCGAAGGCCAGGTCTCGGCCCGCCAGGCCGCG 169
QY 41 IleThrGlyThrTyrTyrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
DB 170 ATCACCGGACCTGCTACCAACAGCTCGGCTCGACCTTCATCGTACCGCGGCGCGAC 229
QY 61 GlyAlaLeuThrGlyThrTyrGlnSerAlaValGlyAsnAlaGlnSerArgTyrValLeu 80
DB 230 GGGCGCCCTAACCGAACCCTACGAGTCGCGCGGACGCGGACGCGCGCTACGCTCG 289
QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
DB 290 ACCGCTGTTACGACAGCGCCCGCGCCACGAGCGGACGCGCGCTCGGTTAGAG 349
QY 101 ValAlaTyrPlyAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGlyGlnTyr 120
DB 350 GTGGCTCGAAGATTAACCTACCGCAACGCCCACTCCGACCAACGCGGAGCGCGCTAC 409
QY 121 ValGlyGlyAlaGlnAlaArgIleAsnThrGlnTyrLeuLeuThrSerGlyThrThrGln 140
DB 410 GTGGCGGCGCGGAGCGGAGATCAACACCGAGTGTCTCTGCTCCGCGACCAACCGAG 469
QY 141 AlaAsnAlaTyrPlySerThrLeuValGlyHisAspThrPheThrIleValIleProSer 160
DB 470 GCCAAGCCTGGAAGTCCAGCGCTGCGCGCCACGACCTTCACCAAGTGAAGCGCTCC 529
QY 161 AlaAlaSerIleAspAlaAlaIleValIleValIleValIleValIleValIleValIle 180
DB 530 GCCGCTTCATCGACGCGCGGAGAAAGCGCGCTCAACACGGAACCGCTCGACGCC 589
QY 181 ValGlnGln 183
DB 590 GTTCAGCAG 598

RESULT 3

US-10-150-762-1
; Sequence 1, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Deaslyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150.762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 638
; TYPE: DNA

ORGANISM: Streptomyces avidinii
US-10-150-762-1

Alignment Scores:

Pred. No.:	1,92e-101	Length:	638
Score:	936.00	Matches:	183
Percent Similarity:	100.004	Conservative:	0
Best Local Similarity:	100.004	Mismatches:	0
Query Match:	100.004	Indels:	0
DB:	14	Gaps:	0

US-09-589-870b-2 (1-183) x US-10-150-762-1 (1-638)

```

QY      1 MetArgLyValIleValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
DB      50 ATCGCGCAAGATCGTCGTGACGATCGCCGTTTCCCTGACCAAGGCTCGATTACGCC 109
QY      21 SerAlaSerAlaAapProSerLyAapSerLyAaGlnValSerAlaAlaGlnAlaGly 40
DB      110 AGCGCTTCGCGACAGCCCTCCAAAGACTCGAAGGCCAGGCTCGCGCGCGAGCGCGC 169
QY      41 IleThrGlyThrTrpTyraEngInLeuGlySerThrHeIleValThrAlaGlyAlaAap 60
DB      170 ATCACCAGCAGCTGTGACACAGCTCGCTCACTTCATCGTACCGCGGCGCGCGAC 229
QY      61 GlyAlaLeuThrGlyThrTyrgLusSerAlaValGlyAaenAlaGlnuSerArgTyraIleu 80
DB      230 GCGGCGCTGACCGGACCTTCAGAGTGGCGGCTCGGCGACCGCGACCGCTACGTCCTG 289
QY      81 ThrGlyAArgTyraAapSerAlaProAlaThrAapGlySerGlyThrAlaLeuGlyTrpThr 100
DB      290 ACCGCTGCTTACGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 349
QY      101 ValAlaIleTrpLyAaenAenTyraArgAaenAlaHisSerAlaThrThrTrpSerGlyGlnTy 120
DB      350 GTGGCTCGAAGAAATACATACCGCAACGCCACTCCGCGCGCGCGCGCGCGCGCGCGCG 409
QY      121 ValGlyGlyValaGlnAlaArgIleAenThrGlnTrpLeuLeuThrSerGlyThrThrGln 140
DB      410 GTGGCGCGCGCGCGCGAGGAGATCAACACCACTGGCTGCTCGCGCGCGCGCGCGCGCG 469
QY      141 AlaAaenAlaTrpLySerThrIleuValGlyHisAapThrHeThrLyValLyProSer 160
DB      470 GCCAAGCGCTGGAAGTCCAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 529
QY      161 AlaAlaSerIleAaenAlaAlaLyValaGlyValAaenGlnLyAaenProLeuAaenAla 180
DB      530 GCGGCTTCATCGACGCGGGAAGGCGCGGCTCAACAAAGCGCGCGCGCGCGCGCGCG 589
QY      181 ValGlnGln 183
DB      590 GTTCAGCAG 598

```

RESULT 4

US-10-244-821-5
Sequence 5, Application US/10244821
Publication NO. US20030143233A1
GENERAL INFORMATION:
APPLICANT: Goshorn, Stephen Charles
APPLICANT: Graves, Scott Stoll
APPLICANT: Schultz, Joanne Elaine
APPLICANT: Lin, Yukang
APPLICANT: Sanderson, James Allen
APPLICANT: Reno, John M.
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
FILE REFERENCE: 690022, 547C3
CURRENT APPLICATION NUMBER: US/10/244, 821
CURRENT FILING DATE: 2002-09-16
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5

LENGTH: 1239
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion
US-10-244-821-5

Alignment Scores:

Pred. No.:	1,53e-91	Length:	1239
Score:	855.50	Matches:	165
Percent Similarity:	98.834	Conservative:	4
Best Local Similarity:	96.494	Mismatches:	1
Query Match:	91.404	Indels:	1
DB:	12	Gaps:	1

US-09-589-870b-2 (1-183) x US-10-244-821-5 (1-1239)

```

QY      14 ThrThrValSerIleThrAla---SerAlaSerAlaAapProSerLyAaenGlyAla 32
DB      724 ACCAGCGTACCGGTGACGTCTGGCTCTGTTGGGACAGACCCCTCCAAAGACTCGAAGGCC 783
QY      33 GlnValSerAlaAlaGlnAlaGlyIleThrGlyThrTrpTyraEngInLeuGlySerThr 52
DB      784 CAGTCTCCGCGCGCGCGCGCGCGCGCATCACCGGCACTGTGTACACACAGCTCGGCTGACC 843
QY      53 PheIleValThrAlaGlyAlaAapGlyAlaLeuThrGlyThrTyrgLusSerAlaValGly 72
DB      844 TTCATGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903
QY      73 AaenAlaGlnuSerArgTyraIleuThrGlyAArgTyraAapSerAlaProAlaThrAapGly 92
DB      904 AACGCGAAGCGCGCTACGTCTGACCGGCTGTTAGACAGCGCGCGCGCGCGCGCGCGCG 963
QY      93 SerGlyThrAlaLeuGlyTrpThrValAlaIleTrpLyAaenAenTyraArgAaenAlaHisSer 112
DB      964 AGCGGACCGCGCGCTCGGTTGAGACGCGGCGCTGGAAGAAATACATACCGCAACGCCCACTCC 1023
QY      113 AlaThrThrTrpSerGlyGlnTyraValaGlyValaGlnAlaArgIleAenThrGlnTrp 132
DB      1024 GGACCGACCTGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1083
QY      133 LeuLeuThrSerGlyThrThrGlnAlaAaenAlaTrpLySerThrLeuValGlyHisAap 152
DB      1084 CTGCTGACCTCCGGGACCGACCGAGCGCAACGCTGGAATCCACGCTGCTCGCGCGCG 1143
QY      153 ThrPheThrLyValIleProSerAlaAlaSerIleAaenAlaAlaLyValaGlyVal 172
DB      1144 ACCTTCACCAAGGTGAAGCGCTCGCGCGCTCATGACGCGGCGGAAGGCGCGCGCTC 1203
QY      173 AaenGlnLyAaenProLeuAaenAlaValaGlnGln 183
DB      1204 AACACGCGCAACCGCTCGACGCGCGCTTCAGCAG 1236

```

RESULT 5

US-10-013-173-5
Sequence 5, Application US/10013173
Publication NO. US20030095977A1
GENERAL INFORMATION:
APPLICANT: Goshorn, Stephen C.
APPLICANT: Graves, Scott Stoll
APPLICANT: Schultz, Joanne Elaine
APPLICANT: Lin, Yukang
APPLICANT: Sanderson, James A.
APPLICANT: Reno, John M.
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
FILE REFERENCE: 690022, 547C1
CURRENT APPLICATION NUMBER: US/10/013, 173
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1239

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion
US-10-013-173-5
```

```

Alignment Scores:
Pred. No.: 1,53e-91 Length: 1239
Score: 855.50 Matches: 165
Percent Similarity: 98.83% Conservative: 4
Best Local Similarity: 96.49% Mismatches: 1
Query Match: 91.40% Indels: 1
DB: 14 Gaps: 1
```

US-09-589-870B-2 (1-183) x US-10-013-173-5 (1-1239)

```

QY 14 ThrThrValSerIleThrTha---SerIaSerAlaAspProSerIlyAspSerIlySAla 32
Db 724 ACCACGGTCACCGTAGCTGTGGCTTGGTGGACAGACCCCTCCAGAGACTCGAAGGCC 783
QY 33 GlnValSerAlaAlaGlnAlaGlyIleThrGlyThrTyrPyrAsnGlnLeuGlySerThr 52
Db 784 CAGGCTCGGCGCCGCGAGCCGCGCATCACCGGCACCTGTACAACAGCTCGGCTCGACC 843
QY 53 PheIleValThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTyrGluSerAlaValGly 72
Db 844 TTCATCGTAGACCGCGGCGCGCGACCGCCCTGACCGAGACCTACGAGTCGGCGCTCGGC 903
QY 73 AsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAspGly 92
Db 904 AACGCCGAGAGCGGCTAGCTCTGACGGATCTGACGACAGAGCGCCCGCGCACCGACGGC 963
QY 93 SerGlyThrAlaLeuGlyTyrThrValAlaTyrPlyAsnAsnTyrArgAsnAlaHisSer 112
Db 964 AGCGGACCGCCCTCGTGTGGACGGTGGCTGGAAGAAATTACTACCGACCCCACTCC 1023
QY 113 AlaThrThrTTPSerGlyGlnTyrValGlyAlaGlyAlaArgIleAsnThrGlnTTP 132
Db 1024 GCGACACAGTGAGGCGGCGCAGTACGTCGCGCGCGCGAGCGAGATCAACACCGAGTGG 1083
QY 133 LeuLeuThrSerGlyThrThrGlnAlaAsnAlaTyrPlySerThrLeuValGlyHisAsp 152
Db 1084 CTGCTGACCTCGGACCAACCGAGGCGCAAGCTTGAAATCCACGCTGTGGCCACGAC 1143
QY 153 ThrPheThrIlyValIlyProSerAlaAlaSerIleAspAlaAlaIlySlyAlaGlyVal 172
Db 1144 ACCTTCACCAAGGTGAAGCCGTCGCGCTCCATCGACGCGCGGAGAGAGCGCGCGTC 1203
QY 173 AsnAsnGlyAsnProLeuAspAlaValGlnGln 183
Db 1204 AACCAACGGCAACCCGCTCGACGCGCGTTCAGCAG 1236
```

RESULT 6

```

US-10-150-762-5
; Sequence 5, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Deastyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1239
```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion
US-10-150-762-5
```

```

Alignment Scores:
Pred. No.: 1,53e-91 Length: 1239
Score: 855.50 Matches: 165
Percent Similarity: 98.83% Conservative: 4
Best Local Similarity: 96.49% Mismatches: 1
Query Match: 91.40% Indels: 1
DB: 14 Gaps: 1
```

US-09-589-870B-2 (1-183) x US-10-150-762-5 (1-1239)

```

QY 14 ThrThrValSerIleThrTha---SerIaSerAlaAspProSerIlyAspSerIlySAla 32
Db 724 ACCACGGTCACCGTAGCTGTGGCTTGGTGGACAGACCCCTCCAGAGACTCGAAGGCC 783
QY 33 GlnValSerAlaAlaGlnAlaGlyIleThrGlyThrTyrPyrAsnGlnLeuGlySerThr 52
Db 784 CAGGCTCGGCGCCGCGAGCCGCGCATCACCGGCACCTGTACAACAGCTCGGCTCGACC 843
QY 53 PheIleValThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTyrGluSerAlaValGly 72
Db 844 TTCATCGTAGACCGCGGCGCGCGACCGCCCTGACCGAGACCTACGAGTCGGCGCTCGGC 903
QY 73 AsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAspGly 92
Db 904 AACGCCGAGAGCGGCTAGCTCTGACGGATCTGACGACAGAGCGCCCGCGCACCGACGGC 963
QY 93 SerGlyThrAlaLeuGlyTyrThrValAlaTyrPlyAsnAsnTyrArgAsnAlaHisSer 112
Db 964 AGCGGACCGCCCTCGTGTGGACGGTGGCTGGAAGAAATTACTACCGACCCCACTCC 1023
QY 113 AlaThrThrTTPSerGlyGlnTyrValGlyAlaGlyAlaArgIleAsnThrGlnTTP 132
Db 1024 GCGACACAGTGAGGCGGCGCAGTACGTCGCGCGCGCGAGCGAGATCAACACCGAGTGG 1083
QY 133 LeuLeuThrSerGlyThrThrGlnAlaAsnAlaTyrPlySerThrLeuValGlyHisAsp 152
Db 1084 CTGCTGACCTCGGACCAACCGAGGCGCAAGCTTGAAATCCACGCTGTGGCCACGAC 1143
QY 153 ThrPheThrIlyValIlyProSerAlaAlaSerIleAspAlaAlaIlySlyAlaGlyVal 172
Db 1144 ACCTTCACCAAGGTGAAGCCGTCGCGCTCCATCGACGCGCGGAGAGAGCGCGCGTC 1203
QY 173 AsnAsnGlyAsnProLeuAspAlaValGlnGln 183
Db 1204 AACCAACGGCAACCCGCTCGACGCGCGTTCAGCAG 1236
```

RESULT 7

```

US-10-244-821-3
; Sequence 3, Application US/10244821
; Publication No. US2003014323A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Deastyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1614
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hunR-LU-10 single chain antibody-genomic streptavidin
; US-10-244-821-3
; US-10-013-173-3

Alignment Scores:
Pred. No.: 2,84e-91 Length: 1614
Score: 854.50 Matches: 166
Percent Similarity: 97.67% Conservative: 2
Best Local Similarity: 96.51% Mismatches: 3
Query Match: 91.29% Indels: 1
DB: 12 Gaps: 1

US-09-589-870b-2 (1-183) x US-10-244-821-3 (1-1614)

QY 12 SerLeuThrThrValSerIleThrAlaSerAlaAspProSerIlyAspSerIlyS 31
Db 1093 ACCTTAGTCACCGTGAGC---TCTGGCTGTGGTTCGGCAGACCCCTCCAAAGGACTCGAAG 1149

QY 32 AlaGlnValSerAlaAlaGlnAlaGlyIleThrGlyThrTyrAsnGlnLeuGlySer 51
Db 1150 GCCCAGGCTCGGCGCCGCGAGCGCGCATCACCGGCACCTGTGTACAAACGACTCGGCTCG 1209

QY 52 ThrPheIleValThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTyrGluSerAlaVal 71
Db 1210 ACCTTCATCTGTACCGCGCGCGCGCGCGCGCTTACCGGAACCTTACGAGTCGGCGCTC 1269

QY 72 GlyAsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAsp 91
Db 1270 GGCAACGGCGAGAGCGCGCTACGTCCTGACCGGCTGTAGACACGCGCGCGCGCACCGAC 1329

QY 92 GlySerGlyThrAlaLeuGlyTyrThrValAlaTyrIlyAsnAsnTyrArgAsnAlaHis 111
Db 1330 GGCAGCGGACCGCCCTCGGTGGACGCTGGCAAGATTAACCGCACCGCCAC 1389

QY 112 SerAlaThrThrTyrSerGlyGlnTyrValGlyAlaGlnAlaArgIleAsnThrGln 131
Db 1390 TCCGCGACCACTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1449

QY 132 TrpLeuLeuThrSerGlyThrThrGlnAlaAsnAlaTyrIlySerThrLeuValGlyHis 151
Db 1450 TGGCTGCTGACCTCGCGGACCAACCGGCGCAACGCTGGAAGTCCACGCTGTCGGCCAC 1509

QY 152 AspThrPheThrIlyValIlySerProSerAlaAlaSerIleAspAlaAlaIlyValGly 171
Db 1510 GACACCTTCACCAAGGTGAAGCGCTCCGCGCTCCATCGACGCGCGCAAGAGCGCGCG 1569

QY 172 ValAsnAsnGlyAsnProLeuAspAlaValGlnGln 183
Db 1570 GTCAACAAAGCGCAACCCGCTCGACGCGCTTCAGCAG 1605

RESULT 8
US-10-013-173-3
; Sequence 3, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1614
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hunR-LU-10 single chain antibody-genomic streptavidin
; US-10-013-173-3
; US-10-150-762-3

Alignment Scores:
Pred. No.: 2,84e-91 Length: 1614
Score: 854.50 Matches: 166
Percent Similarity: 97.67% Conservative: 2
Best Local Similarity: 96.51% Mismatches: 3
Query Match: 91.29% Indels: 1
DB: 14 Gaps: 1

US-09-589-870b-2 (1-183) x US-10-013-173-3 (1-1614)

QY 12 SerLeuThrThrValSerIleThrAlaSerAlaAspProSerIlyAspSerIlyS 31
Db 1093 ACCTTAGTCACCGTGAGC---TCTGGCTGTGGTTCGGCAGACCCCTCCAAAGGACTCGAAG 1149

QY 32 AlaGlnValSerAlaAlaGlnAlaGlyIleThrGlyThrTyrAsnGlnLeuGlySer 51
Db 1150 GCCCAGGCTCGGCGCCGCGAGCGCGCATCACCGGCACCTGTGTACAAACGACTCGGCTCG 1209

QY 52 ThrPheIleValThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTyrGluSerAlaVal 71
Db 1210 ACCTTCATCTGTACCGCGCGCGCGCGCGCGCTTACCGGAACCTTACGAGTCGGCGCTC 1269

QY 72 GlyAsnAlaGluSerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAsp 91
Db 1270 GGCAACGGCGAGAGCGCGCTACGTCCTGACCGGCTGTAGACACGCGCGCGCGCACCGAC 1329

QY 92 GlySerGlyThrAlaLeuGlyTyrThrValAlaTyrIlyAsnAsnTyrArgAsnAlaHis 111
Db 1330 GGCAGCGGACCGCCCTCGGTGGACGCTGGCAAGATTAACCGCACCGCCAC 1389

QY 112 SerAlaThrThrTyrSerGlyGlnTyrValGlyAlaGlnAlaArgIleAsnThrGln 131
Db 1390 TCCGCGACCACTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1449

QY 132 TrpLeuLeuThrSerGlyThrThrGlnAlaAsnAlaTyrIlySerThrLeuValGlyHis 151
Db 1450 TGGCTGCTGACCTCGCGGACCAACCGGCGCAACGCTGGAAGTCCACGCTGTCGGCCAC 1509

QY 152 AspThrPheThrIlyValIlySerProSerAlaAlaSerIleAspAlaAlaIlyValGly 171
Db 1510 GACACCTTCACCAAGGTGAAGCGCTCCGCGCTCCATCGACGCGCGCAAGAGCGCGCG 1569

QY 172 ValAsnAsnGlyAsnProLeuAspAlaValGlnGln 183
Db 1570 GTCAACAAAGCGCAACCCGCTCGACGCGCTTCAGCAG 1605

RESULT 9
US-10-150-762-3
; Sequence 3, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
```

```

; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: humr-LU-10 single chain antibody-genomic streptavidin
US-10-150-762-3

```

Alignment Scores:

```

Pred. No.: 2,84e-91 Length: 1614
Score: 854.50 Matches: 166
Percent Similarity: 97.67% Conservative: 2
Best Local Similarity: 96.51% Mismatches: 3
Query Match: 91.29% Indels: 1
DB: 14 Gaps: 1

```

US-09-589-870b-2 (1-183) x US-10-150-762-3 (1-1614)

```

QY 12 SerLeuThrThrValSerIleThrAlaSerAlaSerAlaAspProSerIysAspSerIys 31
DB 1093 ACCTTAGTACCGCTGAGC---TCTGGCTCTGGTTCCGCGACACCCCTCCAAAGACTCGAAG 1149
QY 32 AlaGlnValSerAlaAlaGlnValGlyIleThrGlyThrTrpTyraSngInleuGlySer 51
DB 1150 GCCGAGGTTCCGCGCGCGCGGCGGCGCATCCCGGACCTGGTACACAGCTCGGCTCG 1209
QY 52 ThrPheIleValThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTyriGluSerAlaVal 71
DB 1210 ACCTTCATCGTACCGCGCGCGCGCGCGCGCGCTGACCGGAACCTACGAGTCCGCGCTCG 1269
QY 72 GlyAsnAlaGluSerArgTyriValLeuThrGlyValArgTyriAspSerAlaProAlaThrAsp 91
DB 1270 GCCAAGCGCGGAGCGCGCTACGTCCTGACCGGCTGTTACGACAGCGCGCGCGCGCAC 1329
QY 92 GlySerGlyThrAlaLeuGlyTrpThrValAlaTrpIleAsnAsnTyriArgAsnAlaHis 111
DB 1330 GCGAGGGGACCGCGCTCGGTGGAGGCGGCGCTGGAAATTAACCTACCGCAACGCCAC 1389
QY 112 SerAlaThrThrTrpSerGlyGlnTyriValGlyAlaGlyAlaArgIleAsnThrGln 131
DB 1390 TCCGGGACCGCGTGGAGCGCGCGAGTACGTCCGCGCGCGCGGAGCGAGATCAACACCCAG 1449
QY 132 TrpLeuLeuThrSerGlyThrThrGlnAlaAsnAlaTrpIleSerThrIleuValGlyHis 151
DB 1450 TGGCTGCTGACCTCCGCGCACACCGCGCGCAACGCTGGAGTCCACGCTGCTCGGCGCAC 1509
QY 152 AspTrpPheThrIleValIysProSerAlaIleAserIleAspAlaAlaIysIysAlaGly 171
DB 1510 GACACCTTCACCAAGGTGAAGCGGTCCGCCCTCCATCGACGCGCGGAGGAGGCGGCG 1569
QY 172 ValAsnAsnGlyAsnProLeuAspAlaValGlnGln 183
DB 1570 GTCAACAACGCGCACCCGCTGACGCGCTTCAGCGAG 1605

```

RESULT 10

```

US/10/244
; Sequence 7, Application US/10244821
; Publication No. US2003014323A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 7
; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion construct
US/10/244,821-7

```

Alignment Scores:

```

Pred. No.: 1.61e-90 Length: 1280
Score: 847.00 Matches: 162
Percent Similarity: 98.79% Conservative: 1
Best Local Similarity: 98.18% Mismatches: 2
Query Match: 90.49% Indels: 0
DB: 12 Gaps: 0

```

US-09-589-870b-2 (1-183) x US/10/244 (1-1280)

```

QY 19 ThrAlaSerAlaSerAlaAspProSerIysAspSerIysAlaGlnValSerAlaAlaGln 38
DB 777 TCTGCTCTGGTTCCGCGACACCCCTCCAAAGACTCGAAGCCCAAGTCTCGGCGCGCGAG 836
QY 39 AlaGlyIleThrGlyThrTrpTyraSngInleuGlySerThrPheIleValThrAlaGly 58
DB 837 GCCGCGATCCCGCGCGCGCGCGCGCGCGCATCCCGGACCTCGGCTGACCTTCATCGTACCGGCGC 896
QY 59 AlaAspGlyAlaLeuThrGlyThrTyriGluSerAlaValGlyAsnAlaGluSerArgTyri 78
DB 897 GCCGAGCGCGCGCTGACCGCGGAACCTACGAGTCCGCGCGCGCGCAACCGCGAGCGCGTAC 956
QY 79 ValLeuThrGlyValArgTyriAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGly 98
DB 957 GTCCGAGCGCGCTGTTACGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCCGT 1016
QY 99 TrpThrValAlaTrpIleAsnAsnTyriArgAsnAlaHisSerAlaThrThrTrpSerGly 118
DB 1017 TGGAGCGGTGCGTGGAAATTAACCTACCGCAACGCCACTCCGAGCACAGTGGAGCGCGC 1076
QY 119 GlnTyriValGlyAlaGlyAlaArgIleAsnThrGlnTrpLeuThrSerGlyThr 138
DB 1077 CAGTACGTCCGCGCGCGCGCGCGCGCGAGATCAACACCAAGTGGCTGCTGACCTCCGCGCAC 1136
QY 139 ThrGlnAlaAsnAlaTrpIleSerThrIleuValGlyHisAspThrPheThrIleValIys 158
DB 1137 ACCGAGCGCGCAACGCTGGAGTCCACGCTGCTCGGCGCACGACCTTCACCAAGGTGAAG 1196
QY 159 ProSerAlaIleAserIleAspAlaAlaIysIysAlaGlyValAsnAsnGlyAsnProLeu 178
DB 1197 CGTCCGCGCGCTCCATCGACGCGCGGAGGAGGCGGCGTCAACAGGCAACCGCGCTC 1236
QY 179 AspAlaValGlnGln 183
DB 1257 GACGCGCTTCAGCGAG 1271

```

RESULT 11

```

US/10/013
; Sequence 7, Application US/10013173
; Publication No. US20030095977A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7

```



```

; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion construct
US/10/013.173-7
```

Alignment Scores:

Pred. No.:	1,61e-90	Length:	1280
Score:	847.00	Matches:	162
Percent Similarity:	98.79%	Conservative:	1
Best Local Similarity:	98.18%	Mismatches:	2
Query Match:	90.49%	Indels:	0
DB:	14	Gaps:	0

US-09-589-870b-2 (1-183) x US/10/013 (1-1280)

```

QY 19 ThrAlaSerAlaSerAlaAapProSerLyAspSerLyAlaGlnValSerAlaAlaGlu 38
   ::::|
Db 777 TCTGGCTGTGTTCCGCGACAGCCCTCCAGAGCTCGAAGGCCAGGCTCGGCCGCGAG 836

QY 39 AlaGlyIleThrGlyThrTrpTyraEngInLeuGlySerThrPheIleValThrAlaGly 58
   |||||
Db 837 GCCGGCATCACCGGCACCTGGTACAAACAGCTCGGCTCGACTTCATCGTACCGCGGGC 896

QY 59 AlaAapGlyAlaLeuThrGlyThrTyrgLusSerAlaValGlyAsnAlaGluSerArgTy 78
   |||||
Db 897 GCCGAGCGGCCCTTGACCGGACCTTACGAGTCCGGCTCGGCAACGCCGAGAGCCGCTAC 956

QY 79 ValLeuThrGlyArGlyTyraSpSerAlaProAlaThrAspGlySerGlyThrAlaLeuGly 98
   |||||
Db 957 GTCCTGACCGGTGTTACGACAGCGCCCGGCCACCGAGCGGACCGGCCGCTCGGT 1016

QY 99 TrpThrValAlaTrpLyAsnAsnTyraGAsnAlaHisSerAlaThrThrTrpSerGly 118
   |||||
Db 1017 TGAGCGGTGGCTCGGAAGATACTACCGCAAGCCGACCTCCGCAACAGTGGACCGGC 1076

QY 119 GlnTyValAlGlyValaGluAlaArgIleAsnThrGlnTrpLeuLeuThrTrpSerGly 138
   |||||
Db 1077 CAGTACGTCCGGCGGCCCGGAGGATCAACACCCAGTGGCTGCTGACCTCCGCGACC 1136

QY 139 ThrGluAlaAsnAlaTrpLySerThrLeuValGlyHisAspThrPheThrLyValLy 158
   |||||
Db 1137 ACCGAGGCCCAACCGCTGGAAGTCCAGCGTGTGGCCACAGACCTTCAACAGGTGAG 1196

QY 159 ProSerAlaAlaSerIleAspAlaAlaLybValaGlyValaAsnGlyAsnProLeu 178
   |||||
Db 1197 CCGTCCGCGCGCTCCATCGACGGGGGAAGAGCCGGGCTCAACACGCAACCGCGCTC 1256

QY 179 AspAlaValaGlnGln 183
   |||||
Db 1257 GAGCGCGTTTCAGCAG 1271
```

RESULT 12

```

US/10/150
; Sequence 7, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150.762
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
```

```

; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: B9E9 single chain antibody-genomic streptavidin fusion construct
US/10/150.762-7
```

Alignment Scores:

Pred. No.:	1,61e-90	Length:	1280
Score:	847.00	Matches:	162
Percent Similarity:	98.79%	Conservative:	1
Best Local Similarity:	98.18%	Mismatches:	2
Query Match:	90.49%	Indels:	0
DB:	14	Gaps:	0

US-09-589-870b-2 (1-183) x US/10/150 (1-1280)

```

QY 19 ThrAlaSerAlaSerAlaAapProSerLyAspSerLyAlaGlnValSerAlaAlaGlu 38
   ::::|
Db 777 TCTGGCTGTGTTCCGCGACAGCCCTCCAGAGCTCGAAGGCCAGGCTCGGCCGCGAG 836

QY 39 AlaGlyIleThrGlyThrTrpTyraEngInLeuGlySerThrPheIleValThrAlaGly 58
   |||||
Db 837 GCCGGCATCACCGGCACCTGGTACAAACAGCTCGGCTCGACTTCATCGTACCGCGGGC 896

QY 59 AlaAapGlyAlaLeuThrGlyThrTyrgLusSerAlaValGlyAsnAlaGluSerArgTy 78
   |||||
Db 897 GCCGAGCGGCCCTTGACCGGACCTTACGAGTCCGGCTCGGCAACGCCGAGAGCCGCTAC 956

QY 79 ValLeuThrGlyArGlyTyraSpSerAlaProAlaThrAspGlySerGlyThrAlaLeuGly 98
   |||||
Db 957 GTCCTGACCGGTGTTACGACAGCGCCCGGCCACCGAGCGGACCGGCCGCTCGGT 1016

QY 99 TrpThrValAlaTrpLyAsnAsnTyraGAsnAlaHisSerAlaThrThrTrpSerGly 118
   |||||
Db 1017 TGAGCGGTGGCTCGGAAGATACTACCGCAAGCCGACCTCCGCAACAGTGGACCGGC 1076

QY 119 GlnTyValAlGlyValaGluAlaArgIleAsnThrGlnTrpLeuLeuThrTrpSerGly 138
   |||||
Db 1077 CAGTACGTCCGGCGGCCCGGAGGATCAACACCCAGTGGCTGCTGACCTCCGCGACC 1136

QY 139 ThrGluAlaAsnAlaTrpLySerThrLeuValGlyHisAspThrPheThrLyValLy 158
   |||||
Db 1137 ACCGAGGCCCAACCGCTGGAAGTCCAGCGTGTGGCCACAGACCTTCAACAGGTGAG 1196

QY 159 ProSerAlaAlaSerIleAspAlaAlaLybValaGlyValaAsnGlyAsnProLeu 178
   |||||
Db 1197 CCGTCCGCGCGCTCCATCGACGGGGGAAGAGCCGGGCTCAACACGCAACCGCGCTC 1256

QY 179 AspAlaValaGlnGln 183
   |||||
Db 1257 GAGCGCGTTTCAGCAG 1271
```

RESULT 13

```

US-10-244-821-87
; Sequence 87, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244.821
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
```

```

; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-244-821-87

Alignment Scores:
Pred. No.: 1.76e-90 Length: 1371
Score: 847.00 Matches: 162
Percent Similarity: 98.79% Conservative: 1
Best Local Similarity: 98.18% Mismatches: 2
Query Match: 90.49% Indels: 0
DB: 12 Gaps: 0

US-09-589-870B-2 (1-183) x US-10-244-821-87 (1-1371)

QY 19 ThrAlaSerAlaSerAlaAspProSerLyAspSerLyAlaGlnValSerAlaAlaGlu 38
Db 874 TCTGGCTGTGTTGGCAGACCCCTCCAGAGACTCGAAGGCCAGGTCTCGCCGCGGAG 933
QY 39 AlaGlyIleThrGlyThrTyrAsnGlnLeuGlySerThrPheIleValThrAlaGly 58
Db 934 GCCGCATCACCGGCACCTGTAACAACAGCTCGCTCGACTTCATCTGTGACCGCGGCG 993
QY 59 AlaAspGlyAlaLeuThrGlyThrTyrGluSerAlaValAlaGlyAsnAlaGluSerArgTyr 78
Db 994 GCCGACGGCGCCCTGACCGGAACCTACGAGTCGGCGCTCGGCAAGCCGAGCGCTAC 1053
QY 79 ValLeuThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGly 98
Db 1054 GTCTGACCGGTGCTTACAGACGCCGCCGCGCACCGAGCGGACCGCCCTCGGT 1113
QY 99 TrpThrValAlaIleTyrLyAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGly 118
Db 1114 TGGACGGTGGCTTGGAAAGATTAACACCGCAAGCCCACTCCGCGACCACTGGAGCGGC 1173
QY 119 GlnTyrValAlaGlyAlaGluAlaArgIleAsnThrGlnTyrLeuLeuThrSerGlyThr 138
Db 1174 CAGTACGTGGCGGCGCCGAGGAGATCAACCCAGTGTGCTGTAACCTCCGCGCAC 1233
QY 139 ThrGluAlaAsnAlaIleTyrLySerThrLeuValGlyHisAspThrPheThrLyValLys 158
Db 1234 ACCGAGGCCAAGCCTGGAAATCCAGCTGTGCGCCACGACCACTTCCAAAGGTGAAG 1293
QY 159 ProSerAlaAlaSerIleAspAlaAlaLysIleValAlaAsnGlyAsnProLeu 178
Db 1294 CCGTCGCGCGCTCATCGACGCGGGAAGAGCGGCGGTCAACAAACGCAACCCGCTC 1353
QY 179 AspAlaValGlnGln 183
Db 1354 GACGCCGTTCAAGCAG 1368

RESULT 14
US-10-244-821-48
; Sequence 48, Application US/10244821
; Publication No. US2003014323A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 1467
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CC49 single chain antibody-genomic streptavidin
US-10-244-821-48

Alignment Scores:
Pred. No.: 1.93e-90 Length: 1467
Score: 847.00 Matches: 162
Percent Similarity: 98.79% Conservative: 1
Best Local Similarity: 98.18% Mismatches: 2
Query Match: 90.49% Indels: 0
DB: 12 Gaps: 0

US-09-589-870B-2 (1-183) x US-10-244-821-48 (1-1467)

QY 19 ThrAlaSerAlaSerAlaAspProSerLyAspSerLyAlaGlnValSerAlaAlaGlu 38
Db 898 TCTGGCTGTGTTGGCAGACCCCTCCAGAGACTCGAAGGCCAGGTCTCGCCGCGGAG 957
QY 39 AlaGlyIleThrGlyThrTyrAsnGlnLeuGlySerThrPheIleValThrAlaGly 58
Db 958 GCCGCATCACCGGCACCTGTAACAACAGCTCGCTCGACTTCATCTGTGACCGCGGCG 1017
QY 59 AlaAspGlyAlaLeuThrGlyThrTyrGluSerAlaValAlaGlyAsnAlaGluSerArgTyr 78
Db 1018 GCCGACGGCGCCCTGACCGGAACCTACGAGTCGGCGCTCGGCAAGCCGAGCGCTAC 1077
QY 79 ValLeuThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGly 98
Db 1078 GTCTGACCGGTGCTTACAGACGCCGCCGCGCACCGAGCGGACCGCCCTCGGT 1137
QY 99 TrpThrValAlaIleTyrLyAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyrSerGly 118
Db 1138 TGGACGGTGGCTTGGAAAGATTAACACCGCAAGCCCACTCCGCGACCACTGGAGCGGC 1197
QY 119 GlnTyrValAlaGlyAlaGluAlaArgIleAsnThrGlnTyrLeuLeuThrSerGlyThr 138
Db 1198 CAGTACGTGGCGGCGCCGAGGAGATCAACCCAGTGTGCTGTAACCTCCGCGCAC 1257
QY 139 ThrGluAlaAsnAlaIleTyrLySerThrLeuValGlyHisAspThrPheThrLyValLys 158
Db 1258 ACCGAGGCCAAGCCTGGAAATCCAGCTGTGCGCCACGACCACTTCCAAAGGTGAAG 1317
QY 159 ProSerAlaAlaSerIleAspAlaAlaLysIleValAlaAsnGlyAsnProLeu 178
Db 1318 CCGTCGCGCGCTCATCGACGCGGGAAGAGCGGCGGTCAACAAACGCAACCCGCTC 1377
QY 179 AspAlaValGlnGln 183
Db 1378 GACGCCGTTCAAGCAG 1392

RESULT 15
US-10-013-173-48
; Sequence 48, Application US/10013173
; Publication No. US2003009597A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C1
; CURRENT APPLICATION NUMBER: US/10/013,173
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 1467
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: CC49 single chain antibody-genomic streptavidin
; OTHER INFORMATION: fusion sequence
US-10-013-173-48

```

Alignment Scores:

Pred. No.:	1,936-90	length:	1467
Score:	847.00	Matches:	162
Percent Similarity:	96.79%	Conservative:	1
Best Local Similarity:	99.18%	Mismatches:	2
Query Match:	90.49%	Indels:	0
DB:	14	Gaps:	0

US-09-589-870B-2 (1-183) x US-10-013-173-48 (1-1467)

QY	19	ThrlAlaSerAlaSerAlaAepProSerItyrApsSerIyAlaGlnAlaGlu	38
Db	898	TTGGCTTGGTTGGCGAGACCCCTCCAGAGCTCGAAGGCCAGGTCTGGCGCCAG	957
QY	39	AlaGlyIleThrGlyThrTrpTyraSngInIleuGlySerThrPheIleValThrlaGly	58
Db	958	GCCGGACATCACCGGACCTGGTATACACAGCTCGGCTGACCTTCATGTGTACCGGGGC	1017
QY	59	AlaAepGlyAlaLeuThrGlyThrTyGlnSerAlaValGlyAsnAlaGlnSerArgTyr	78
Db	1018	GCCGACGGCGCCCTGACCGGAACCTACAGATCGGGCGTCGGCAAGCCGGAAGCGCGTAC	1077
QY	79	ValIleuThrGlyArgTyrAepSerAlaProAlaThrApsGlySerGlyThrlAlaLeuGly	98
Db	1078	GTCTGTACCGGTCGTATACAGACGCGCCCGGCACCGACGGCAGGGGACCGCCCTCGGT	1137
QY	99	TrpThrValAlaTrpIyAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGly	118
Db	1138	TGGACGGTGGCTGGAGAAATPACTACCGCAACGCCACTCGGACACCACTGGAGCGGC	1197
QY	119	GlnTyrValaGlyGlyAlaGlnAlaArgIleAsnThrGlnTrpLeuIleuThrSerGlyThr	138
Db	1198	CAGTACGTGGCGGGCGCCGAGGCGAGATCAACACCAGCTGGCTGTACTCCGGCACCC	1257
QY	139	ThrGlnAlaAsnAlaTrpIySerThrLeuValGlyHisAepThrPheThrIyValIyIs	158
Db	1258	ACCGAGGCCAACGCTGGAAGTCCACGCTGGTCGGCCACGACACTTCACCAAGGTGAAG	1317
QY	159	ProSerAlaAlaSerIleAepAlaAlaIyAlaIyAlaGlyValaAsnAemGlyAsnProIeu	178
Db	1318	CCGTCTCCGCGCTTCATCGACGCGCGAAGAGCCGCGCTCAACAACGCGAACCCGCTC	1377
QY	179	AspAlaValGlnGln	183
Db	1378	GACGCGCTTCAGCAG	1392

Search completed: October 28, 2003, 03:00:31
Job time : 259 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: October 27, 2003, 21:58:22 ; Search time 61 Seconds
(without alignments)
1324.149 Million cell updates/sec

Title: US-09-589-870B-2

Perfect score: 936
Sequence: 1 MKRIVAAIVASLTVSITR.....IDAKKAGVNGNPLDAVQO 183

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=x1h
-Q=/cgnt2.1/USPTO.epool/US0589870/runac.27102003.104447.16637/app.query.fasta_1.327
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=opt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0589870.0CGN1.1.56.0runac.27102003.104447.16637 -NCPu=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -OS=BL0CK=100 -LONLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgnt2.6/prodata/2/ina/5A_COMB.seq.*
2: /cgnt2.6/prodata/2/ina/5B_COMB.seq.*
3: /cgnt2.6/prodata/2/ina/6A_COMB.seq.*
4: /cgnt2.6/prodata/2/ina/6B_COMB.seq.*
5: /cgnt2.6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgnt2.6/prodata/2/ina/backfillseq1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	100.0	552	5	PCT-US93-05240-13
2	936	100.0	638	4	US-08-831-399-1
3	936	100.0	638	4	US-09-381-330-1
4	936	100.0	638	4	US-09-366-862-1
5	936	100.0	638	4	US-09-368-772-1
6	936	100.0	1131	6	5168049-1
7	846	90.4	1135	2	US-08-491-988-4
8	837.5	89.5	525	1	US-07-924-028A-2
9	837.5	79.6	1286	2	US-08-491-988-6
10	687.5	73.5	1173	4	US-09-142-974B-4
11	682	72.9	1176	4	US-09-142-974B-3
12	681.5	72.8	1257	2	US-08-491-988-8

13	675	72.1	384	4	US-08-831-399-15	Sequence 15, Appl
14	675	72.1	384	4	US-09-366-862-15	Sequence 15, Appl
15	675	72.1	384	4	US-09-368-772-15	Sequence 15, Appl
16	675	72.1	387	1	US-08-211-833-1	Sequence 1, Appl
17	675	72.1	387	1	US-08-434-718-1	Sequence 1, Appl
18	675	72.1	387	1	US-07-780-717C-6	Sequence 6, Appl
19	145	15.5	604	4	US-08-831-399-3	Sequence 3, Appl
20	145	15.5	604	4	US-09-366-862-3	Sequence 3, Appl
21	145	15.5	604	4	US-09-368-772-3	Sequence 3, Appl
22	142.5	15.2	484	1	US-08-554-586-1	Sequence 1, Appl
23	110.5	11.8	3381	4	US-09-336-447A-6	Sequence 6, Appl
24	105.5	11.3	3018	4	US-09-206-942-40	Sequence 4, Appl
25	105.5	11.3	3036	4	US-09-206-942-38	Sequence 38, Appl
26	105.5	11.3	3129	4	US-09-252-991A-13873	Sequence 13873, A
27	105.5	11.3	4131	4	US-09-252-991A-13873	Sequence 13873, A
28	105.5	11.3	8211	4	US-09-252-991A-13856	Sequence 13856, A
29	101.5	10.8	3300	2	US-08-194-290-6	Sequence 6, Appl
30	101.5	10.8	3300	2	US-08-614-377A-6	Sequence 6, Appl
31	101.5	10.8	3300	3	US-09-142-648B-6	Sequence 6, Appl
32	99.5	10.6	3108	4	US-09-206-942-72	Sequence 72, Appl
33	99.5	10.6	4937	1	US-08-038-682-3	Sequence 3, Appl
34	99.5	10.6	4937	1	US-08-302-832-3	Sequence 3, Appl
35	99.5	10.6	4937	2	US-08-530-198-3	Sequence 3, Appl
36	99.5	10.6	4937	2	US-08-469-880-3	Sequence 3, Appl
37	99.5	10.6	4937	2	US-08-728-470-3	Sequence 3, Appl
38	99.5	10.6	4937	2	US-08-617-697-3	Sequence 3, Appl
39	99.5	10.6	4937	2	US-08-719-641-3	Sequence 3, Appl
40	99.5	10.6	4937	4	US-09-206-942-70	Sequence 70, Appl
41	99.5	10.6	9323	1	US-08-038-682-6	Sequence 6, Appl
42	99.5	10.6	9323	1	US-08-302-832-6	Sequence 6, Appl
43	99.5	10.6	9323	2	US-08-530-198-6	Sequence 6, Appl
44	99.5	10.6	9323	2	US-08-469-880-6	Sequence 6, Appl
45	99.5	10.6	9323	2	US-08-728-470-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
PCT-US93-05240-13
; Sequence 13, Application PC/TUS9305240
; GENERAL INFORMATION:
; APPLICANT: NAGARAJAN, VASANTHA
; TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN FROM BACILLUS
; TITLE OF INVENTION: SUBTILLIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DU PONT COMPANY
; STREET: BARLEY MILL PLAZA 36
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19880-0036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05240
; FILING DATE: 19930527
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GEIGER, KATHLEEN W
; REFERENCE/DOCKET NUMBER: CR 9029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-1118
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
PCT-US93-05240-13

Alignment Scores:
Pred. No.: 3,02e-91 Length: 552
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-589-870B-2 (1-183) x PCT-US93-05240-13 (1-552)

```

QY      1 MetArgLysIleValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
Db      1 ATGCGCAGATCGTGTGACACCATTCGGCGTTTCCCTGACACGAGTCTGATACGGCC 60
QY      21 SerAlaSerAlaAspProSerIleAspSerIleValSerAlaAlaGluAlaGly 40
Db      61 AGCGCTTGGCAGACCCCTCCAGAGACTCGAAGGCCCGAGTCTCGCGCCGAGGCCGCGC 120
QY      41 IleThrGlyThrTrpTyrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
Db      121 ATCACCAGCAGCTGTGACACAGCTCGGCTGACCTTCATCTGATACCGCGGCGCCGAC 180
QY      61 GlyAlaLeuThrGlyThrTyrGluSerAlaValGlyAsnAlaGluSerArgTyrValLeu 80
Db      181 GGCGCCCTGACCGGAACCTACAGAGTCGGCGCTGGCAAGCCGAGAGCGGCTACGTCCTG 240
QY      81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThr 100
Db      241 ACCGGTGTGAGACAGCGCCCGCCGACCGAGCGAGCGGACCGCCCTCGGTGGAGC 300
QY      101 ValAlaTrpLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyr 120
Db      301 GTGGCTTGGAGAAATTAATACCGCAACGCCCACTCCGACACCACTGAGCGGCGCAATAC 360
QY      121 ValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuThrSerGlyThrThrGlu 140
Db      361 GTCCGCGCGCGCGAGCGGATCAACACCGAGTGGCTGCTGACCTCCGCGGACCGCAG 420
QY      141 AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSer 160
Db      421 GCCAAGCGCTGGAATTCACAGCTGTGCGCCACGACCTTCACCAAGGTGAAGCGCTCC 480
QY      161 AlaAlaSerIleAspAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
Db      481 GCCGCTTCATGACGCGCGGAGAAAGCGCGGCTCAACACAGCGCAACCCGCTCGACGCC 540
QY      181 ValGlnGln 183
Db      541 GTTCAGCAG 549

```

RESULT 2

US-08-831-399-1
Sequence 1, Application US/08831399
Patent No. 6312916

GENERAL INFORMATION:

APPLICANT: Kopeckzki, Erhard; Muller, Rainer;
APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
TITLE OF INVENTION: Recombinant Inactive Core
TITLE OF INVENTION: Streptavidin Mutants
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2

```

/ OPERATING SYSTEM: PC-DOS
/ SOFTWARE: WordPerfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/831,399
/ FILING DATE: 1-April-1997
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: DE 196 13 053.0
/ FILING DATE: 1-April-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: DE 196 37 718.8
/ FILING DATE: 16-September-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hanson, No. 6312916man D.
/ REGISTRATION NUMBER: 30,946
/ REFERENCE/DOCKET NUMBER: HUBR 1105
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 688-9200
/ TELEFAX: (212) 838-3884
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 638 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 50..598
/ OTHER INFORMATION:
/ OTHER INFORMATION: peptide, and 122..598 to mat peptide.

```

US-08-831-399-1

Alignment Scores:
Pred. No.: 3.62e-91 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-589-870B-2 (1-183) x US-08-831-399-1 (1-638)

```

QY      1 MetArgLysIleValAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20
Db      50 ATGCGCAGATCGTGTGACACCATTCGGCGTTTCCCTGACACGAGTCTGATACGGCC 109
QY      21 SerAlaSerAlaAspProSerIleAspSerIleValSerAlaAlaGluAlaGly 40
Db      110 AGCGCTTGGCAGACCCCTCCAGAGACTCGAAGGCCCGAGTCTCGCGCCGAGGCCGCGC 169
QY      41 IleThrGlyThrTrpTyrAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60
Db      170 ATCACCAGCAGCTGTGACACAGCTCGGCTGACCTTCATCTGATACCGCGGCGCCGAC 229
QY      61 GlyAlaLeuThrGlyThrTyrGluSerAlaValGlyAsnAlaGluSerArgTyrValLeu 80
Db      230 GGCGCCCTGACCGGAACCTACAGAGTCGGCGCTTCGCGCAAGCGGCTACGTCCTG 289
QY      81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThr 100
Db      290 ACCGGTGTGAGACAGCGCCCGCCGACCGAGCGGAGCGGCGGCTCGGTGGAGC 349
QY      101 ValAlaTrpLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTyr 120
Db      350 GTGGCTTGGAGAAATTAATACCGCAACGCCCACTCCGACACCACTGAGCGGCGCAGTAC 409
QY      121 ValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuThrSerGlyThrThrGlu 140
Db      410 GTCCGCGCGCGCGAGCGGATCAACACCGAGTGGCTGCTGACCTCCGCGACCAACCGAG 469
QY      141 AlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLysValLysProSer 160
Db      470 GCCAAGCGCTGGAATTCACAGCTGTGCGCCACGACCTTCACCAAGGTGAAGCGCTCC 529

```

OY	161	AlAlaAserILeSpbAlaAlaLysLysAlaGlyValAsnAsnGlyAsnProLeuAspAla	160
Db	530	GGCGCTCCATCGACGCGGAGAAAGCCGGCGTCAACAACGGCAACCCGCTCGAGCC	589
OY	181	ValGlnGln 183	
Db	590	GTTCAAGCAG 598	
RESULT 3			
US-09-381-430-1			
Sequence 1, Application US/09381430			
Patent No. 6368813			
GENERAL INFORMATION:			
APPLICANT: Reznik, Gabriel O.			
APPLICANT: Sano, Takeshi			
APPLICANT: Vajda, Sandor			
APPLICANT: Cantor, Charles			
TITLE OF INVENTION: MULTIFLAVOR STREPTAVIDIN			
FILE REFERENCE: 1586-50152			
CURRENT APPLICATION NUMBER: US/09/381,430			
CURRENT FILING DATE: 2000-03-23			
PRIOR APPLICATION NUMBER: PCT/US98/04931			
PRIOR FILING DATE: 1998-03-13			
PRIOR APPLICATION NUMBER: 60/040,771			
PRIOR FILING DATE: 1997-03-14			
NUMBER OF SEQ ID NOS: 2			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO: 1			
LENGTH: 638			
TYPE: DNA			
ORGANISM: Streptomyces avidinii1			
US-09-381-430-1			
Alignment Scores:			
Pred. No.: 3,69e-91 Length: 638			
Score: 936.00 Matches: 183			
Percent Similarity: 100.00% Conservative: 0			
Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 100.00% Indels: 0			
DB: 4 Gaps: 0			
US-09-589-870B-2 (1-183) x US-09-381-430-1 (1-638)			
OY	1	MetatgLyLeIeValValAlaAlaAlaIleAlaValSerLeuThrThrValSerIleThrAla 20	
Db	50	ATGCCCAAGATGTCGTCGACGACATCGCGCTTCCCTGACCAACGGCTCGATTACGGCC 109	
OY	21	SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGluAlaGly 40	
Db	110	AGCGCTTCGGCAGACCCCTCCAAAGCATGGAAGGCCAGATCTCGGCGCGCGAGCGCGC 169	
OY	41	IleThrGlyThrTrpTyraAsnGlnLeuGlySerThrPheIleValThrAlaGlyAlaAsp 60	
Db	170	ATCACCAGGACCTCGTACCAACAGCTCGGCTGCACCTTATCTGTACCGCGGCGCGAC 229	
OY	61	GlyAlaLeuThrGlyThrTyrgLysSerAlaValaGlyAsnAlaGluSerArgTyValLeu 80	
Db	230	GGCGCCCTGACCGGAACCTACAGATCGCGCGTGGCAACGCGAGAGCGGCTACGTCCTG 289	
OY	81	ThrgLyArgTyraAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTTPThr 100	
Db	290	ACCGTCTTACGACAGCGCCCGCGCCACCGACGAGCGGCGACCGCCCTCGGTGGACG 349	
OY	101	ValAlaTrpLysAsnAsnTyraArgAsnAlaHisSerAlaThrThrTrpSerGlyGlnTy 120	
Db	350	GTCGCTCGGAAGAAATTAACGCAACGCCCACTCGGACGCCAGCGGCGGCGAGTAC 409	
OY	121	ValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThrGlu 140	
Db	410	GTCGCGGGCGGCGAGCGGATCAACACCCAGTGGCTGTGACCTTCGGCACACCGAG 469	

```

QY      141  ALaAaNaLaTbPLySSeTThrLeuVaIGlYHIsAspThPhetNryLsVallyBProSer 160
Db      470  GCCAAGCGCTTGAAGTCCACCGCTGCGGCCACGACACTTCCACCAAGTAAGCCGTC 529
QY      161  ALaAaSerIleAspAlaAlaLYeLaYLaGlyValaNaenGIYAsnProLeuAspAla 180
Db      530  GCCCGCTTCATCGACGGCGCAAGAAGCCGGCTCAACAACGGCAACCGCTCGACGCC 589
QY      181  ValGInGIn 183
Db      590  GTTCAGCAG 598

RESULT 4
US-09-366-862-1
; Sequence 1, Application US/09366862
; Patent No. 6391571
; GENERAL INFORMATION:
; APPLICANT: Kopeczki, Erhard; Muller, Rainier;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Feife & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/366,862
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,399
; FILING DATE: 1-April-1997
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6391571man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..598
; OTHER INFORMATION: Positions 50..121 correspond to sig peptide, and
; OTHER INFORMATION: 122..598 to mat peptide.
US-09-366-862-1

Alignment Scores:
Pred. No.:      3,69e-91      Length:      638
Score:          936.00      Matches:      183
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              Gaps:      0

```

```

US-09-589-870B-2 (1-183) x US-09-366-862-1 (1-638)
QY 1 MetArgLysIleValAlaIleAlaIleValSerLeuThrThrValSerIleThrAla 20
Db 50 ATGCCAGATGCTGCTTCCAGCCATCGCCGTTTCCCTACCAACGAGTTCGATTACGGCC 109
QY 21 SerAlaSerAlaAspProSerIlyAspSerIlyAspIleValSerAlaAlaGluAglly 40
Db 110 AGCGCTTCGGCAGACCCCTCCCAAGAGCTGGAAGGCCAGAGTCTCGGCCGCCAGGCCG 169
QY 41 IleThrGlyThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 60
Db 170 ATCACCAGGACCTGCTACCAACAGCTCGGCTCGACCTTATGCTGACCGCGGCGCCGAC 229
QY 61 GlYAlaLeuThrGlyThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 80
Db 230 GCGCGCTTCGACCGGAACCTACAGAGTGGCGCTCGGCAACGCCGAGAGCCGCTACGTC 289
QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
Db 290 ACCGGTCGTTACGACAGCGCCCGGCCACGAGCGGACCGGCGCCCTCGATTGAGAGC 349
QY 101 ValAlaIleTyrPlyAsnAsnTyrArgAsnAlaHisSerAlaThrThrThrSerGlyTyr 120
Db 350 GTGGCTTGGAAGAAATTAATACTACGCAACGCCCACTCGGCAACCGTGGAGCGCCAGATC 409
QY 121 ValGlyGlyAlaGlyAlaAlaArgIleAsnThrGlnTyrPleuLeuThrSerGlyThrThrGlu 140
Db 410 GTCCGCGCGCGCGGAGCGGAGATCAACACCCAGTGGCTGCTACCTCCGGCACCCAGAG 469
QY 141 AlaAsnAlaIleTyrPlySerThrLeuValGlyHisAspThrPheThrIlyValIlyProSer 160
Db 470 GCCAAGCGCTGGAAGATCCACGCTGTGGCGCCAGCACCTTACCAAGGTGAAGCCGCTCC 529
QY 161 AlaAlaSerIleAspAlaAlaIlyLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
Db 530 GCCGCTTCATGACGCGGCGGAAGAGCGGCGTCAACACGGAACCGCTCGACGCC 589
QY 181 ValGlnGln 183
Db 590 GTTCAGCAG 598

RESULT 5
US-09-368-772-1
; Sequence 1, Application US/09368772
; Patent No. 6417331
; GENERAL INFORMATION:
; APPLICANT: Kopeckki, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/368,772
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,399
; FILING DATE: 1-April-1997
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6417331man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..598
; OTHER INFORMATION: Positions 50..121 correspond to sig peptide, and
; OTHER INFORMATION: 122..598 to mat peptide.
; US-09-368-772-1

Alignment Scores:
Pred. No.: 3,696-91 Length: 638
Score: 936.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-589-870B-2 (1-183) x US-09-368-772-1 (1-638)
QY 1 MetArgLysIleValAlaIleAlaIleValSerLeuThrThrValSerIleThrAla 20
Db 50 ATGCCAGATGCTGCTTCCAGCCATCGCCGTTTCCCTACCAACGAGTTCGATTACGGCC 109
QY 21 SerAlaSerAlaAspProSerIlyAspSerIlyAspIleValSerAlaAlaGluAglly 40
Db 110 AGCGCTTCGGCAGACCCCTCCCAAGAGCTGGAAGGCCAGTCTCGCGCGCGAGCCGCGC 169
QY 41 IleThrGlyThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 60
Db 170 ATCACCAGGACCTGCTACCAACAGCTCGGCTCGACCTTATGCTGACCGCGGCGCCGAC 229
QY 61 GlYAlaLeuThrGlyThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 80
Db 230 GCGCGCTTCGACCGGAACCTACAGAGTGGCGCTCGGCAACGCCGAGAGCCGCTACGTC 289
QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTyrThr 100
Db 290 ACCGGTCGTTACGACAGCGCCCGGCCACGAGCGGACCGGACCCCTCGATTGAGAGC 349
QY 101 ValAlaIleTyrPlyAsnAsnTyrArgAsnAlaHisSerAlaThrThrThrSerGlyTyr 120
Db 350 GTGGCTTGGAAGAAATTAATACTACGCAACGCCCACTCGGCAACCGTGGAGCGCCAGATC 409
QY 121 ValGlyGlyAlaGlyAlaAlaArgIleAsnThrGlnTyrPleuLeuThrSerGlyThrThrGlu 140
Db 410 GTCCGCGCGCGGAGCGGAGATCAACACCCAGTGGCTGCTACCTCCGGCACCCAGAG 469
QY 141 AlaAsnAlaIleTyrPlySerThrLeuValGlyHisAspThrPheThrIlyValIlyProSer 160
Db 470 GCCAAGCGCTGGAAGATCCACGCTGTGGCGCCAGCACCTTACCAAGGTGAAGCCGCTCC 529
QY 161 AlaAlaSerIleAspAlaAlaIlyLysAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
Db 530 GCCGCTTCATGACGCGGCGGAAGAGCGGCGTCAACACGGAACCGCTCGACGCC 589
QY 181 ValGlnGln 183
Db 590 GTTCAGCAG 598

RESULT 6

```


5168049-1
 Patent No. 5168049
 APPLICANT: MEADE, HARRY M.; GARWIN, JEFFREY L.
 TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN-LIKE
 POLYPEPTIDES
 NUMBER OF SEQUENCES: 6
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/185,329
 FILING DATE: 21-APR-1988
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 656,873
 FILING DATE: 02-OCT-1984
 SEQ ID NO: 1:
 LENGTH: 1131
 5168049-1

Alignment Scores:
 Pred. No.: 8,176-91 Length: 1131
 Score: 936.00 Matches: 183
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-589-870B-2 (1-183) x 5168049-1 (1-1131)

QY 1 MetArgLysIleValIleAlaIleAlaIleSerLeuThrThyValSerIleThrAla 20
 DB 480 ATGGCGAAGATCGTCCTGGAGCATCGCGCTTCCCTGACCGCTCGATTACGGCC 539
 QY 21 SerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGluIle 40
 DB 540 AGCGCTTCGGACACCCCTTCAGAGACTCGAAGGCCAGGCTTCGGCGCCGAGCGCCG 599
 QY 41 IleThrGlyThrTrpTyrAsnGlnLeuGlySerThrPheIleValThrAlaGluIle 60
 DB 600 ATCACCAGGACCTGGTACACACAGCTCGGCTGACCTTATCGTGAACGGCGCGCCGAC 659
 QY 61 GlyAlaLeuThrGlyThrTyrGluSerAlaValIleGlyAsnAlaGluSerArgTyrVal 80
 DB 660 GGGGCGCTGACCGAAGCTAGAGTGGCGCTGGGACGCGCAGAGCGCTACGCTCG 719
 QY 81 ThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrpThr 100
 DB 720 ACCGCTCGTACACACAGCGCCCGCCGACCGAGCGGACCGCCCTCGCTGAGCG 779
 QY 101 ValAlaIleTrpLysAsnAsnTyrArgAsnAlaIleSerAlaThrThrTrpSerGlyGlnTyr 120
 DB 780 GTGGCTGGAGAAATACTACCGCAACGCCCACTCCGCGACACGCGAGCGCCAGTAC 839
 QY 121 ValIleGlyValAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThrGlu 140
 DB 840 GTGCGGGGCGCCGAGGAGGAGTCAACACCCACTGCTGCTCGCCGACACCGAG 899
 QY 141 AlaAsnAlaTrpLysSerThrLeuValGlyLysAspThrPheThrLysValLysProSer 160
 DB 900 GCCAAGCGCTGAAAGTCCAGCTGGTGGCCGACGACACCTTCCACCAAGGTAAAGCCGTC 959
 QY 161 AlaIleSerIleAsnAlaIleAlaLysValAlaGlyValAsnAsnGlyAsnProLeuAspAla 180
 DB 960 GCCGCTCCATCGACCGGGAAGAGCGCGGCTCAACACGCGAACCGCTCGACGCG 1019
 QY 181 ValIleGln 183
 DB 1020 GTTCACGAG 1028

RESULT 7
 US-08-491-988-4
 Sequence 4, Application US/08491988
 Patent No. 597316
 GENERAL INFORMATION:
 APPLICANT: EPENETOS, AGAMENON A.
 APPLICANT: SPOONER, ROBERT A.

APPLICANT: DEONARAIN, MAHENDRA
 TITLE OF INVENTION: Compounds for targeting
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MACAULAY NISSEN GOLDBERG KIEL & HAND, LLP
 STREET: 261 MADISON AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016-2391
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/491,988
 FILING DATE: 18-DEC-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: GOLDBERG, JULES E.
 REGISTRATION NUMBER: 24,408
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-986-4090
 TELEFAX: 212-818-9479
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1356 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 40..1344
 US-08-491-988-4

Alignment Scores:
 Pred. No.: 4,176-81 Length: 1356
 Score: 846.00 Matches: 162
 Percent Similarity: 99.39% Conservative: 1
 Best Local Similarity: 98.78% Mismatches: 1
 Query Match: 90.38% Indels: 0
 DB: Gaps: 0

US-09-589-870B-2 (1-183) x US-08-491-988-4 (1-1356)

QY 20 AlaSerAlaSerAlaAspProSerLysAspSerLysAlaGlnValSerAlaAlaGluIle 39
 DB 853 GCTGGCGGACCTGGACAGACCCGTCAGAGCTCCAAAGCTCAAGCTTCTGACCCGAAAGCT 912
 QY 40 GlyIleThrGlyThrTrpTyrAsnGlnLeuGlySerThrPheIleValThrAlaGluIle 59
 DB 913 GGTATCACCTGCGACCTCGTATACCACTGGGGTGCATCTTATCTGACCCCTGGTGG 972
 QY 60 AspGlyAlaLeuThrGlyThrTyrGluSerAlaValIleGlyAsnAlaGluSerArgTyrVal 79
 DB 973 GACGAGGCTGACCTGCGCACCTAGCAATCTGGTTGGTGAACGAAATCCCGCTACGTA 1032
 QY 80 LeuThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAlaLeuGlyTrp 99
 DB 1033 CTGACTGGCGCTTATGACTCTGACCTGCGACCGAGTGGCTGTGGTCCCTCGGGCTGG 1092
 QY 100 ThrValAlaIleTrpLysAsnAsnTyrArgAsnAlaIleSerAlaThrThrTrpSerGlyGln 119
 DB 1093 ACTGTGGCTGGAAAAACAACATATGTAATGGCACAGGCCACATACGTGCTGGCCAA 1152
 QY 120 TyrValIleGlyValAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSerGlyThrThr 139
 DB 1153 TACGTTGGCGGCTGAGGCTCGTATCAACCTCAGTGGCTGTAAACATCCCGCACCTACC 1212
 QY 140 GluAlaAsnAlaIleTrpLysSerThrLeuValGlyLysAspThrPheThrLysValLysPro 159

Db 1213 GAAGGATGATGGAATCGACCTAGTAGTCACTGACACCTTACCAAACTTAAGCCT 1272
Qy 160 SerAlaAlaSerIleAspAlaAlaLysIleValAsnAsnGlyAsnProLeuasp 179
Db 1273 TCTGCTGATGATTGATGCTGCCAAGAAAGCAGTAAACACGGTAAACCTCTAGAC 1332
Qy 180 AlaValGlnGln 183
Db 1333 GCTGTTCAAGCAA 1344
RESULT 8
US-07-924-028A-2
; Sequence 2, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz, Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 40 05 874
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBER 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-924-028A-2
Alignment Scores:
Pred. No.: 9 01e-81 length: 525
Score: 837.50 Matches: 164
Percent Similarity: 95.93% Conservative: 1
Best Local Similarity: 95.35% Mismatches: 6
Query Match: 89.48% Indels: 1
DB: 1 Gaps: 1
US-09-589-870B-2 (1-183) x US-07-924-028A-2 (1-525)
Qy 12 SerLeuThrThrValSerIleThrAlaSerAlaAspProSerIleLysAspSerIle 31
Db 1 TCTAGAACTAGTAGATCCATC---GAGGGTAGGTCTATGACCCGTCGAAGACTCCAAA 57
Qy 32 AlaGlnValSerAlaAlaGlnValGlyIleThrGlyThrTyrTyrAsnGlnLeuGlySer 51

Db 58 GCTCAGATTCTGACAGCCGAAGCTGATATCACTGGCACCTGGTATACCAACTGGGGTGC 117
Qy 52 ThrPheIleValIleThrAlaGlyAlaAspGlyAlaLeuThrGlyThrTyrGlySerAlaVal 71
Db 118 ACTTTCATTGTGACCGCTGCTGGCGGACCGAGCTGATGCTGACCTAGCAATCTCGGATT 177
Qy 72 GlyAsnAlaGlySerArgTyrValLeuThrGlyArgTyrAspSerAlaProAlaThrAsp 91
Db 178 GGTAAACGAGATCCCGCTAGCTAGTACTGACCGCGTTATGACTGTGCACCTGCCACCAT 237
Qy 92 GlySerGlyThrAlaLeuGlyTyrThrValAlaTatPlyAsnAsnTyrArgAsnAlaHis 111
Db 238 GGCTCTGATACCGCTGCTGGCTGACCTGTGCTGGAACAACTATGTAATGCGAC 297
Qy 112 SerAlaThrThrTyrPseGlyGlnTyrValGlyAlaGlnAlaArgIleAsnThrGln 131
Db 298 ACGCCACTAGTGTGCTGCGCAATACCTTGGCGGTGTGAGGCTCGATACACACTCAG 357
Qy 132 TyrLeuLeuThrSerGlyThrThrGlnAlaAsnAlaTyrPlySerThrLeuValGlyHis 151
Db 358 TGGCTGTTAAACATCCGACACTAACCGAAGCATGCAATCGACACTAGTAGTCAT 417
Qy 152 AspThrPheThrIleValIlePseSerAlaIleSerIleAspAlaAlaLysIleValGly 171
Db 418 GACACCTTTACCAAGTTAAGCTTTCTGCTGTACATTGATGCTGCCAAGAAAGCAGGC 477
Qy 172 ValAsnAsnGlyAsnProLeuAspAlaValGlnGln 183
Db 478 GTAAACAAACGGTAAACCTCTAGACCGCTGTTACGAA 513
RESULT 9
US-08-491-988-6
; Sequence 6, Application US/08491988
; Patent No. 5973116
; GENERAL INFORMATION:
; APPLICANT: EPERNETOS, AGAMENON A.
; APPLICANT: SPOONER, ROBERT A.
; APPLICANT: DEONARAIN, MAHENDRA
; TITLE OF INVENTION: Compounds for targeting
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCANLAY NISSEN GOLDBERG KIEL & HAND, LLP
; STREET: 261 MADISON AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/491,988
; FILING DATE: 18-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDBERG, JULES E.
; REGISTRATION NUMBER: 24,408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-986-4090
; TELEFAX: 212-818-9479
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 40..1284

```

US-08-491-988-6
Alignment Scores:
Pred. No.: 2,328-70 Length: 1296
Score: 745.00 Matches: 142
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 98.61% Mismatches: 1
Query Match: 79.59% Indels: 0
DB: 2 Gaps: 0

US-09-589-870B-2 (1-183) x US-08-491-988-6 (1-1296)
QY 20 AAsESeAlAsESeAlAsESeSerLyAsESeLyAlAGInValSeSeAlAGlAla 39
DB 853 GCTGGCGGACCTCGACAGACCCGTCAGAGACTCAAGACTCAAGATTCTGCAAGCGAAGCT 912
QY 40 GYlIeThrGlyThrTTrPTrAsnGlnLeuGlySeThrPheIleValThrAlaGlyVala 59
DB 913 GGTATCACTGGGACCTGGTATACCAACTGGGGGTGACTTTTCATTGTACCGCTGGTGG 972
QY 60 AepGlyAlaLeuThrGlyThrTyGluSerAlaValGlyAsnAlaGluSerArgTyVal 79
DB 973 GACGAGCTCTGACCTGCGACCCATCCAACTCTGGGTGTGAACGCAAGATCCCGTACCTTA 1032
QY 80 LeuThrGlyArgTyAspSerAlaProAlaThrAspGlySeGlyThrAlaLeuGlyTrp 99
DB 1033 CTGACTGGCCGTTATGACTCTGCACCTCGCACCGATGGCTCTGTGATCCGCTCTGGCGTGG 1092
QY 100 ThrValAlaTrpLyAsnAsnTyArgAsnAlaHisSerAlaThrThrTrpSerGlyGln 119
DB 1093 ACTGGCTTGGAAAAACACATCATCTGTAATGGGACACAGCGCACATCAGTGGTGGCCAA 1152
QY 120 TyrValGlyGlyAlaGluAlaArgIleAsnThrGlnTrpLeuThrSerGlyThrThr 139
DB 1153 TACGTGGCGGGCTGAGAGCTCGTATCAACACTCATGTGGCTGTTAACATCCGGCACTTACC 1212
QY 140 GluAlaAsnAlaTrpLySeSerThrLeuValGlyHisAspThrPheThrTyrValAlaAspPro 159
DB 1213 GAAGGATGCATGAGAAATCCACACTAGTAGTCTGACACCTTATCCAAAGTTAAGCT 1272
QY 160 SerAlaAlaSer 163
DB 1273 TCTGCTGCTAGC 1284

RESULT 10
US-09-142-974B-4
Sequence 4, Application US/09142974B
Patent No. 6451995
GENERAL INFORMATION:
APPLICANT: Cheung, Nai-Kong V.
APPLICANT: Larson, Steven M.
APPLICANT: Guo, Hong-Fen
APPLICANT: Rivlin, Ken
APPLICANT: Sadelain, Michel
TITLE OF INVENTION: Single Chain Fv Constructs of Anti-Ganglioside GD2
TITLE OF INVENTION: Antibodies
FILE REFERENCE: MSK.P-013-USNP
CURRENT APPLICATION NUMBER: US/09/142,974B
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: PCT/US97/04427
PRIOR FILING DATE: 1997-03-20
PRIOR APPLICATION NUMBER: 60/013,703
PRIOR FILING DATE: 1996-03-20
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 1173
TYPE: DNA
ORGANISM: Murine
FEATURE:
OTHER INFORMATION: 366-bcfv-8teptavidin
US-09-142-974B-4

```

Alignment Scores:					
Pred. No.:	2,74e-64	Length:	1173		
Score:	687.50	Matches:	133		
Percent Similarity:	92.52%	Conservative:	4		
Best Local Similarity:	90.48%	Mismatches:	3		
Query Match:	73.45%	Indels:	7		
DB:	4	Gaps:	1		
US-09-589-870B-2 (1-183) x US-09-142-974B-4 (1-1173)					
QY	17 SerIeThrAlaSerAlaSerAlaAspProSerLyAspSerLyValAGlnValSerAla 36				
Db	697 TCAgTCAcCGCTCTCTCAAGCGCC-----GCTGAATCGGtGCT 735				
QY	37 AlaGlnAlaGlyIleThrGlyThrTyrAsnGlnLeuGlySerThrPheIleValThr 56				
Db	736 GCTGAACAGATATCAACCGGACCTGTATACACAGCTCGGCTGACTTCATGTGACC 795				
QY	57 AlaGlyAlaAspGlyAlaLeuThrGlyThrTyrGlnSerAlaValGlyAsnAlaGlnSer 76				
Db	796 GCGGGCCCGCAAGCGCCCTGACCGGAACCTACAGTGGCCGCTGGCAACGCCGAGAGC 855				
QY	77 ArgTyrAlaLeuThrGlyArgTyrAspSerAlaProAlaThrAspGlySerGlyThrAla 96				
Db	856 CGCTACGCTCTGACCGGTCGTTACGACAGCGCCCGGCCACCGAGCGAGCGACCGCC 915				
QY	97 LeuGlyTrpThrValAlaIleTyrLysAsnAsnTyrArgAsnAlaHisSerAlaThrThrTyr 116				
Db	916 CTGGTTGGACGGTGGCTGTGAAGAAATTAACCGCAACCCCACTCCGGAGCACGTGG 975				
QY	117 SerGlyGlnTyrValGlyGlyAlaGlnAlaArgIleAsnThrGlnTyrPleuLeuThrSer 136				
Db	976 AGCGCCAGTACGTCGGCGCGCGCGAGCGAGCATCAACCCCGTGGCTGTGACCTCC 1035				
QY	137 GlyThrThrGlnAlaAsnAlaIleTyrLysSerThrLeuValGlyHisAspThrPheThrLys 156				
Db	1036 GGCAACAACGAGGCCAAGCGCTGGAAGTCCACGCTGTGGCGCACGACCTTCACCAAG 1095				
QY	157 ValLysProSerAlaAlaSer 163				
Db	1096 GTGAAGCGTCTCGCGCTCC 1116				
RESULT 11					
US-09-142-974B-3					
: Sequence 3, Application US/09142974B					
: Patent No. 6451995					
GENERAL INFORMATION:					
APPLICANT: Cheung, Nai-Kong V.					
APPLICANT: Larson, Steven M.					
APPLICANT: Guo, Hong-Fen					
APPLICANT: Rivlin, Ken					
APPLICANT: Sadelain, Michel					
TITLE OF INVENTION: Single Chain FV Constructs of Anti-Ganglioside GD2					
TITLE OF INVENTION: Antibodies					
FILE REFERENCE: MSK-P-013-USNP					
CURRENT APPLICATION NUMBER: US/09/142,974B					
CURRENT FILING DATE: 1998-09-18					
PRIOR APPLICATION NUMBER: PCT/US97/04427					
PRIOR FILING DATE: 1997-03-20					
PRIOR APPLICATION NUMBER: 60/013,703					
PRIOR FILING DATE: 1996-03-20					
NUMBER OF SEQ ID NOS: 5					
SOFTWARE: Patent Ver. 2.1					
SEQ ID NO 3					
LENGTH: 1176					
TYPE: DNA					
ORGANISM: Murine					
FEATURE:					
OTHER INFORMATION: 5f11-scfv-streptavidin					
NAME/KEY: unsure					
LOCATION: (37)					
NAME/KEY: unsure					
LOCATION: (79)					

COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/831,399
FILING DATE: 1-April-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 196 13 053.0
FILING DATE: 1-April-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 196 37 718.8
FILING DATE: 16-September-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6312916man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: HUBR 1105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..384
US-08-831-399-15

Alignment Scores:
Pred. No.: 1,256-63 Length: 384
Score: 675.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.12% Indels: 0
DB: 4 Gaps: 0

US-09-589-870b-2 (1-183) x US-08-831-399-15 (1-384)

QY 37 AAGUAAGAGIYIETHRGYTHRTPTYRAENGINLEUGYSETRHPHEILEVALTHR 56
DB 4 GCCGAAGCTGATACATGCGACCTGCTATACCACTGGCGGCGACTTTCATTGGAC 63
QY 57 AAGCAGAAAGAGIYALALALALALALALALALALALALALALALALALALAL 76
DB 64 GCTGGCTGACGAGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 123
QY 77 ARGTYVALLEUHRGTYARGTYRAESPSEALAPROALATHRAPGLYSEGTTHRALA 96
DB 124 CGCTACCTACTGACGCGCTGATGACTGACCTGACCTGACCTGACCTGACCTG 183
QY 97 LEUGLYTRPHTHVALALATRPYASASANTYRARGASNAIAHISSEALATHTHTTP 116
DB 184 CTGGGCTGAGCTGGCTTGGAACAACTATCTGATATCGCACACTGCTGCTAGCTG 243
QY 117 SERGIYGINTYVALIGIYIAGIYALARGILEASNTHTGINTRPLEUHTHRSER 136
DB 244 TCTGGCCATATACGTTGGCGGCTGAGGCTGATATCACTGACCTGCTGTTACATCC 303
QY 137 GATGTHRGILUALAANAATRPYASERTHLEUVALGHTHSAAPTHTPHTHTLV 156
DB 304 GGCACTACCGAAGGAATGATGAATAATGACACTAGTAGCATGACCTTTACCAA 363
QY 157 VALYEPROSEALAAASER 163
DB 364 GTTAAGCTTCTGCTGCTAGC 384

RESULT 14
US-09-366-862-15
Sequence 15, Application US/09366862
Patent No. 6391571

GENERAL INFORMATION:
APPLICANT: Kopecki, Erhard; Muller, Rainer;
APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
TITLE OF INVENTION: Recombinant Inactive Core Streptavidin Mutants
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/366,862
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/831,399
FILING DATE: 1-April-1997
APPLICATION NUMBER: DE 196 13 053.0
FILING DATE: 1-April-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 196 37 718.8
FILING DATE: 16-September-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6312917man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: HUBR 1105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..384
US-09-366-862-15

Alignment Scores:
Pred. No.: 1,256-63 Length: 384
Score: 675.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.12% Indels: 0
DB: 4 Gaps: 0

US-09-589-870b-2 (1-183) x US-09-366-862-15 (1-384)

QY 37 AAGUAAGAGIYIETHRGYTHRTPTYRAENGINLEUGYSETRHPHEILEVALTHR 56
DB 4 GCCGAAGCTGATACATGCGACCTGCTATACCACTGGCGGCTGACCTTTCATTGGAC 63
QY 57 AAGCAGAAAGAGIYALALALALALALALALALALALALALALALALALALAL 76
DB 64 GCTGGCTGACGAGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 123
QY 77 ARGTYVALLEUHRGTYARGTYRAESPSEALAPROALATHRAPGLYSEGTTHRALA 96
DB 124 CGCTACCTACTGACGCGCTGATGACTGACCTGACCTGACCTGACCTGACCTG 183
QY 97 LEUGLYTRPHTHVALALATRPYASASANTYRARGASNAIAHISSEALATHTHTTP 116
DB 184 CTGGGCTGAGCTGGCTTGGAACAACTATCTGATATCGCACACTGCTGCTAGCTG 243

QY 117 SerGIYrValIleuThrGlyValAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSer 136
| | | | |
Db 244 TCTGCCAATACGTTGGCGGCTGCTGATCAACACCTGAGCTGTGTAACATCC 303
| | | | |
QY 137 GLYThrThrGluAlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLys 156
| | | | |
Db 304 GGCACCTACCGAAGCGAATGCATGGAATGCACATAGTAGTCATGACACCTTTACCAA 363
| | | | |
QY 157 ValLysProSerAlaAlaSer 163
| | | | |
Db 364 GTTAAGCCTTCTGCTGCTAGC 384
| | | | |

RESULT 15

US-09-368-772-15
; Sequence 15, Application US/09368772
; Patent No. 6417331
; GENERAL INFORMATION:
; APPLICANT: Kopeckzi, Erhard; Muller, Rainer;
; APPLICANT: Engh, Richard; Schmitt, Urban; Deger, Arno; Brandstetter, Hans
; TITLE OF INVENTION: Recombinant Inactive Core Streptavidin Mutants
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 Kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09368,772
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,399
; FILING DATE: 1-April-1997
; APPLICATION NUMBER: DE 196 13 053.0
; FILING DATE: 1-April-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 196 37 718.8
; FILING DATE: 16-September-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6417331man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..384
US-09-368-772-15

Alignment Scores:

Pred. No.: 1.25e-63 Length: 384
Score: 675.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.12% Indels: 0
DB: 4 Gaps: 0

US-09-589-870b-2 (1-183) x US-09-368-772-15 (1-384)

QY 37 AlaGluAlaGlyIleThrGlyThrTrpTyraSerGlnLeuGlySerThrPheIleValThr 56
| | | | |
Db 4 GCCGAAGCTGGATATCACACGACACCTGGATATACCAACAGGGGTGACTTTCATTGTGACC 63
| | | | |
QY 57 AlaGluAlaAspGlyAlaLeuThrGlyThrTyrgLysSerAlaValAlaGlyAsnAlaGlySer 76
| | | | |
Db 64 GCTGCTGTACGAGAGCTCTGACTGACCTACGAAATCTGGCGTTGGTAACGCAATCC 123
| | | | |
QY 77 ArgTyrValIleuThrGlyArgTyraSerAlaProAlaThrAspGlySerGlyThrAla 96
| | | | |
Db 124 CGTAAGTACTGACTGGCGGCTTATGACTCTGACCTGCCACCGAATGCGCTGTGTAACCGCT 183
| | | | |
QY 97 LeuGlyTrpThrValAlaIleTrpLysAsnAsnTyraArgAsnAlaHisSerAlaThrThr 116
| | | | |
Db 184 CTGGCTGTGACTGTGGCTTGAACAAACACTATGTAATGGCACAGTCCACTACGTGG 243
| | | | |
QY 117 SerGIYrValIleuThrGlyValAlaGluAlaArgIleAsnThrGlnTrpLeuLeuThrSer 136
| | | | |
Db 244 TCTGCCAATACGTTGGCGGCTGCTGATCAACACCTGAGCTGTGTAACATCC 303
| | | | |
QY 137 GLYThrThrGluAlaAsnAlaTrpLysSerThrLeuValGlyHisAspThrPheThrLys 156
| | | | |
Db 304 GGCACCTACCGAAGCGAATGCATGGAATGCACACTAGTAGTCAATGACACCTTTACCAA 363
| | | | |
QY 157 ValLysProSerAlaAlaSer 163
| | | | |
Db 364 GTTAAGCCTTCTGCTGCTAGC 384
| | | | |

Search completed: October 28, 2003, 01:31:56
Job time : 66 secs